



## **On Optimal Forecasting with Soft Computation for Nonlinear Time Series**

SHUENN-REN CHEN

*Department of Business Administration, Kaohsiung 833, Kaohsiung, ChengShiu Institute of Business & Technology, Taiwan*

BERLIN WU<sup>1</sup>

berlin@math.nccu.edu.tw

*Department of Mathematical Sciences, Wen-shan 116, Taipei, National Chengchi University, Taiwan*

**Abstract.** In this paper we present a new approach on optimal forecasting by using the fuzzy set theory and soft computing methods for the dynamic data analysis. This research is based on the concepts of fuzzy membership function as well as the natural selection of evolution theory. Some discussions in the sensitivity of the design of fuzzy processing will be provided. Through the design of genetic evolution, the *AIC* criteria is used as the adjust function, and the fuzzy memberships function of each gene model are calculated. Simulation and empirical examples show that our proposed forecasting technique can give an optimal forecasting in time series analysis.

**Keywords:** optimal forecasting, soft computation, genetic modeling, membership function

### **1. Introduction**

One weakness of many proposed models for time series clearly rests on the assumption of having no structure changes during the whole dynamic process. In the real case, it is very difficult to construct an appropriate model, which can honestly explain the trend of an underlying time series, such as exchange rate or index of stock volume. Two fundamental questions that often arise are: (1) Does there exist an appropriate statistical model that can account for this underlying process? (2) Does the dynamic process agree with a single linear or nonlinear equation? (Need we use more than one equation, e.g. threshold model, to fit the time series?)

If a priori model family is correctly chosen, procedures of model construction such as parameter estimation, diagnosis and forecasting will make sense. But if the underlying time series demonstrates certain structure changes, it is natural to detect those change points or change periods before modeling the whole process. Otherwise, using the traditional techniques for model construction, we may hardly get an appropriated *good model* from a model-base to justify the nonlinear time series.

In the nonlinear time series analysis, literatures have proposed a lot of detecting and testing methods about change points before model construction. While in the real case, it often demonstrated that a time series did change regime gradually. The change points have illustrated senses of fuzziness and heredity. Many patterns of change structure exhibit certain kinds of duration, those phenomena should not be treated as a mere sudden turning at a fixed time.

In many expositions of fuzzy methods, fuzzy techniques are described as an alternative to a more traditional statistical approach. The concept of fuzzy sets (logic) provides a more realistic and moderate approach, by referring to fuzzy measure and classification concept human brain utilizes in dynamic surroundings, to handle the phenomenon of multi-complexities and uncertainties. Because fuzzy theory has intrinsic features of linguistic variables, it can minimize trouble on dealing with uncertain problems. Therefore, fuzzy theory has been widely applied in many fields such as aerospace, mechanical engineering, medical science, power generation, and geology, etc. Among these fields, the application of fuzzy control systems is even more popular; see Nguyen and Sugeno (1998).

The problem of change regime detection in a time series has been examined by many researchers. For instance, Sastri, Flores and Valdes (1989) presented a study of performance comparison for six time-series change detection procedures. Inclan and Tiao (1994) proposed an iterative procedure to detect variance changes based on a centered version of the cumulative sums of squares presented by Brown, Durbin and Evans (1975). Rukhin (1997) studied the classical change-point estimation problem in the bayesian setting. Wu and Chen (1999) demonstrated fuzzy statistical methods in change periods detection. Recently, Wu and Chung (2002) use genetic algorithms to parameters ( $d$ ,  $r$ ) estimation for threshold autoregressive models.

However, those detecting techniques are based on the assumption that the underlying time series conducts a significant change point characteristic. Nevertheless, we must indicate that in dealing with the time series with switching regimes, we had better consider the change period's property instead of the change points detection. Because many patterns of change structure in time series exhibit a certain kind of duration, those phenomena should not be treated as a mere sudden turning at a certain time, cf. Wu and Chen (1998). For instance: (i) the exchange rate may go up or down gradually after a new financial policy performs. (ii) a national monetary supply of  $M_1A$  or  $M_2$  may change their trend at different period of time according to the national economic conditions. In fact, the semantics of the term "change point" is vague or uncertain (interesting readers may refer to any popular dictionary such as Webster's New Dictionary).

Fuzzy logic is ideal for modeling and controlling complex, nonlinear systems because it systematically handles ambiguity. In contrast to the black-and white, true-and -false world of traditional logic, fuzzy logic lets you use degrees of truth to describe system behavior. This dramatically simplifies the solution of many engineering and decision support problems, even when rigorous mathematical models are not available.

In this research, we make use of an integrated procedure to construct a genetic model for an optimal forecasting for dynamic data. Firstly, we choose certain initial gene models from a built model-base and decide the generation from the data. Then, under the adjust function we calculate the degree of membership for each gene models in each generation. After deciding the memberships of the fitness, we examine the evolution trend and choose the *leading model* from all generation. Finally, we decide the change periods by checking whether other candidate models have replaced the leading model. The modeling process we perform is called *genetic modeling*. This detecting method for change periods will be called *genetic detecting*. The simulation and empirical results showed that our technique is an efficient and realistic procedure in *forecasting* a nonlinear time series. Especially, when

the change is gradual, our model construction procedure demonstrated a better data explanation as well as prediction.

## 2. What is Genetic Modeling?

The Genetic Algorithm (GA) proposed by Holland (1970) is based on a derivative-free stochastic optimization method on the concepts of natural selection and evolutionary processes. Evolution theory stressed the fact that the existence of all living things is in accordance with the rule of *survival of the fittest*. Darwin suggested that new breeds or classes of living things come into existence through the processes of reproduction, cross-over, and mutation among existing organisms. The concepts of the evolution theory have been translated into algorithms to search for solutions to problems in a more *natural* way.

In the following material we show how the evolutionary concepts and natural selection are used in the nonlinear time series analysis. Applying to the classical evolution theory into the analysis of real data, we may look the dynamic processes of a nonlinear time series as a result of heredity of certain gene models. Hence if we partition the whole time series into  $N$  generation, and examine each generation with degree of fitness we may capture some heredity property through evolution.

The procedure of our method is by choosing initial gene models from a model-base and examines the heredity property through the generations. Since the ARIMA model family demonstrates a very good property in explaining the short-term stochastic process, we will choose five basic ARMA models: AR(1), AR(2), MA(1), MA(2) and ARMA(1,1) as our initial gene models. The well-known AIC criterion is used as our adaptive measurement for the gene models, where  $AIC = n \ln \hat{\sigma}_\varepsilon^2 + 2(p + q)$ . That is the best appropriate model chosen according to the AIC criteria.

### 2.1. Membership of Fitness

For each generation, in order to measure the degree of fitness for gene models, we need to give a definition about degree of fitness. The membership of fitness is a useful tool, which can help us find a leading model through the evolution.

#### Definition 2.1 Membership of fitness for gene models

Let  $\{C_1, C_2, \dots, C_k\}$  be the  $k$  gene models. Let  $A_{ij}$  be the AIC value of the  $j^{\text{th}}$  ( $j = 1, 2, \dots, k$ ) gene model in the  $i^{\text{th}}$  ( $i = 1, 2, \dots, N$ ) generation and  $A_i = \min_{1 \leq j \leq k} \{A_{ij}\}$ .

Then the membership of fitness for each gene model written by  $M_{ij}$  is defined as

$$M_{ij} = 1 - \frac{A_{ij} - A_i}{\sum_{j=1}^k (A_{ij} - A_i)}. \quad (2.1)$$

*Example 2.1.* Let  $X_t$  follow an AR (1) process with  $X_t = 5 + 0.5X_{t-1} + \varepsilon_t$ ;  $\varepsilon_t \sim N(0,1)$ . With a size of 250 simulated data, we separated it into 6 generations. Suppose we

Table 2.1. AIC values for 5 candidate models in 6 generation.

Models	AR(1)	AR(2)	MA(1)	MA(2)	ARIMA(1 0 1)
Generation 1	23.99	26.44	27.81	28.03	26.44
Generation 2	5.08	7.59	7.77	9.65	7.56
Generation 3	1.91	4.50	2.44	4.56	4.79
Generation 4	14.30	16.83	15.62	18.16	16.77
Generation 5	1.91	4.79	1.83	4.71	3.06
Generation 6	-1.80	1.02	-0.33	0.60	1.03

choose 5 initial gene models AR(1), AR(2), MA(1), MA(2) and ARMA(1,1) from the ARMA(p,q) family. The AIC values for 5 candidate models in 6 generations are illustrated at Table 2.1.

We calculate the AIC value for each gene model and then apply equation (2.1) to compute the membership of gene models for the first generation. Since

$$\begin{aligned} \sum_{j=1}^5 (A_{1j} - A_1) &= (23.99 - 23.99) + (26.44 - 23.99) + (27.81 - 23.99) \\ &\quad + (28.03 - 23.99) + (26.44 - 23.99) \\ &= 12.76. \end{aligned}$$

$$M_{11} = 1 - \frac{A_{11} - A_1}{\sum_{j=1}^5 A_{1j} - A_1} = 1 - \frac{23.99 - 23.99}{12.76} = 1,$$

$$M_{12} = 1 - \frac{A_{12} - A_1}{\sum_{j=1}^5 A_{1j} - A_1} = 1 - \frac{26.44 - 23.99}{12.76} = 0.81,$$

$$M_{13} = 1 - \frac{A_{13} - A_1}{\sum_{j=1}^5 A_{1j} - A_1} = 1 - \frac{27.81 - 23.99}{12.76} = 0.70,$$

$$M_{14} = 1 - \frac{A_{14} - A_1}{\sum_{j=1}^5 A_{1j} - A_1} = 1 - \frac{28.03 - 23.99}{12.76} = 0.68,$$

$$M_{15} = 1 - \frac{A_{15} - A_1}{\sum_{j=1}^5 A_{1j} - A_1} = 1 - \frac{26.44 - 23.99}{12.76} = 0.81.$$

Table 2.2 shows the result for the membership of best-fitted model for all six generations.

From Table 2.2, we can find that the smaller the AIC value is, the larger the membership is. That is, the fittest model will account for the simulated data fairly. Therefore we will choose the highest membership function as our best-fitted model.

Table 2.2. Memberships for 5 candidate models in 6 generation.

Gene Models	AR(1)	AR(2)	MA(1)	MA(2)	ARMA(1,1)
Generation 1	1.00	0.81	0.70	0.68	0.81
Generation 2	1.00	0.80	0.78	0.63	0.80
Generation 3	1.00	0.00	0.94	0.69	0.67
Generation 4	1.00	0.95	0.97	0.62	0.96
Generation 5	0.99	0.59	1.00	0.60	0.83
Generation 6	1.00	0.70	0.85	0.75	0.70

## 2.2. The Leading Model

After calculating the membership of fitness for gene models at each generation, we are going to find the leading model, which dominates the underlying dynamic process.

Traditionally, we use *AIC* criterion as our decision rule to choose the most appropriate model for the underlying time series. While in this research, we select the most appropriate models not only by the minimum *AIC* value but also those models that close to the minimum *AIC* under a certain significance level  $\alpha$ . That is, the most appropriate model may not be unique, all models which are close to the minimum *AIC* value are selected as the most appropriate models.

If we examine the most appropriate models through generations, we may find certain regulations among them. One way is to look at the outcome number of the most appropriated for each gene model. If the frequency of the most appropriate model reaches a confidence level  $\lambda$ ,  $0.5 < \lambda < 1$ , then we can say that this gene model can account for the time series fairly. Definition 2.2 and 2.3 give the precise definitions for the above idea. Under the significant  $\alpha$ -level, we claim that a leading model is one that takes the highest frequency for the most appropriated in all generations.

### Definition 2.2 Most appropriate models under the significant $\alpha$ -level

Let  $A(i)$  be the fuzzy set of gene models at the  $i^{\text{th}}$  generation. Under the significant level  $\alpha$  ( $0 < \alpha < 1$ ), the most appropriate models  $A(i)$  at the generation  $i$  are defined by  $A_\alpha(i) = \{A(x) \mid A(i; x) \geq 1 - \alpha, x \in U\}$ .

*Example 2.2.* From Table 2.1 of Example 2.1 the fuzzy set of  $A_\alpha(5)$  at the 5<sup>th</sup> generation can be written as:

$$A_\alpha(5) = \frac{0.99}{AR(1)} + \frac{0.59}{AR(2)} + \frac{1}{MA(1)} + \frac{0.60}{MA(2)} + \frac{0.83}{ARMA(1,1)}.$$

Under the significance level  $\alpha = 0.1$ , we get  $A_{0.1}(5) = \{C_1, C_3\}$ . Hence, at the 5<sup>th</sup> generation, there are 2 most appropriate models: AR(1) and MA(1).

Table 2.3. The  $n_i(j)$  values for 6 generation in Example 2.1.

Gene Models	AR(1)	AR(2)	MA(1)	MA(2)	ARMA(1,1)
Generation 1	1	0	0	0	0
Generation 2	1	0	0	0	0
Generation 3	1	0	1	0	0
Generation 4	1	1	1	0	1
Generation 5	1	0	1	0	0
Generation 6	1	0	0	0	0
$S(j)$	1	1/6	3/6	0	1/6

In order to choose a leading model for a nonlinear time series, we propose the following definition 2.3, where  $n_i(j) = 1$  if the  $j^{\text{th}}$  gene model belongs to  $A_{\alpha}$ , and the  $n_i(j) = 0$  otherwise.

**Definition 2.3** *The leading model*

Let  $\{X_t, t = 1, 2, \dots, m\}$  be a time series,  $C = \{C_j : j = 1, 2, \dots, k\}$  be the set of gene models and  $S(j) = \frac{1}{N} \sum_{i=1}^N n_i(j)$  (where  $N$  is the number of generation), be the frequency of the gene model  $C_j$  being chosen as the most appropriate model at each generation. Given a confidence level  $\lambda$  ( $0.5 < \lambda < 1$ ), if  $\max\{S(j), j = 1, \dots, k\} = S(l) \geq \lambda$ , then we say that the  $l^{\text{th}}$  gene model is the leading model for time series  $\{X_t\}$ .

*Example 2.3.* From Table 2.1 of Example 2.1, we find the AR(1) model get the maximum memberships at the 5 gene models in the first generation. By Definition 2.2, under the significant level  $\alpha = 0.1$ , we have  $A_{0.1}(j) = 0.9$ . Hence  $n_1(1) = 1, n_1(2) = n_1(3) = n_1(4) = n_1(5) = 0$ .

Similarly, the best-fitted models at the 5<sup>th</sup> generation are AR(1) and MA(1). Hence denote  $n_5(1) = n_5(3) = 1$  and  $n_5(2) = n_5(4) = n_5(5) = 0$ . Similarly, we can get the Table 2.3.

Obviously, if we choose the confidence level  $\lambda = 0.9$ , then  $S(1) = 1 > 0.9$ , it indicates that the AR(1) is the leading model for this time series.

### 2.3. An Integrated Genetic Modeling Process

By observing the evolution of time series with  $N$  generations, we can find the memberships of candidate models as well as the leading model. On the other hand if there exist certain change points/periods for the leading model, we may also say that the underlying time series exists a structure change.

In this section, we propose a decision for the change period identification. The main idea is that if the membership of leading model at the  $i^{\text{th}}$  generation is less than  $A_{\alpha}$ , it means that the last generation of best-fitted model has transferred to other gene models. Naturally the underlying time series encounters a structure change. And we can say that the change period occurred at this generation.

The following procedure demonstrates the identification and decision steps.

#### 2.4. Algorithm for Identification of Change Periods of a Time Series

- (1) Take the set of the gene models  $C = \{C_j : j = 1, 2, \dots, k\}$  and separate the time series into  $k$  generations.
- (2) Choose an adaptive function, and calculate the membership of fitness  $M_{ij}$  for each gene model  $j$  and each generation  $i$ .
- (3) Under the significance level  $\alpha$  and the fuzzy set  $A_\alpha$  ( $i = 1, 2, \dots, N$ ), deciding the  $n_i(j)$ ,  $n_i(j) = 1$  if  $C_j$  belongs to the most appropriate model  $A_\alpha(i)$  at the  $i^{\text{th}}$  generation, otherwise  $n_i(j) = 0$ .
- (4) Calculating the frequency of  $S(j)$ . Under the confidence level  $\lambda \in (0,1)$ , if there exist  $j$  such that  $S(j) \geq \lambda$ , then the leading model is assigned to the  $j^{\text{th}}$  gene model, go to step 6; otherwise go to step 5.
- (5) The time series exhibits a very unstable process, it maybe a random walk, a chaos or encounter too many intervention from outside system.
- (6) Stop.

#### 2.5. The Optimal Forecasting

After identifying the leading model, we can do the forecasting job for this nonlinear time series.

Let  $f_{1,t}, f_{2,t}, \dots, f_{k,t}$  be the forecasting values after  $t$  steps from the gene models  $C_1, C_2, \dots, C_k$  respectively. We consider that the conditional expectation of the variable being forecasted is a linear combination of the available forecasts. Thus, when combining the individual forecasts  $f_{1,t}, f_{2,t}, \dots, f_{k,t}$ , the optimal forecasting  $F_t$  is deducted according to equation (2.2)

$$F_t = w_1 f_{1,t} + w_2 f_{2,t} + \dots + w_k f_{k,t}, \quad \text{where} \quad w_j = \frac{S(j)}{\sum_{j=1}^k S(j)}, \quad j = 1, \dots, k. \quad (2.2)$$

Clearly, this forecasting by weighted memberships combination in general will be more reasonable and efficient than the combined forecasting of equal weight.

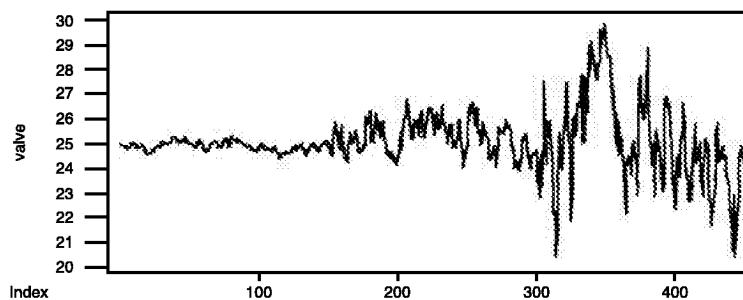


Figure 3.1. Trend for the time series model (3.1) (mean = 25.61 SD = 1.26).

### 3. Simulated Examples

The time series data with size 450 was generated from three models, see model (3.1), (3.2) and (3.3). An AR(1) process with three different noise terms is shown in Figure 3.1. An ARCH(1) process with three different variances is shown in Figure 3.2. In Figure 3.3, an ARMA(1,2) process with three different moving average noises terms is shown.

$$X_t = \begin{cases} 5 + 0.8X_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, 0.1) & \text{if } 1 \leq t \leq 150 \\ 5 + 0.8X_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, 0.45) & \text{if } 151 \leq t \leq 300 \\ 5 + 0.8X_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, 1.25) & \text{if } 301 \leq t \leq 450 \end{cases} \quad (3.1)$$

$$X_t = \begin{cases} \sigma_t \varepsilon_t, \sigma_t^2 = 5 + 0.5X_{t-1}^2, \varepsilon_t \sim N(0, 1) & \text{if } 1 \leq t \leq 150 \\ \sigma_t \varepsilon_t, \sigma_t^2 = 5 + 0.2X_{t-1}^2, \varepsilon_t \sim N(0, 1) & \text{if } 151 \leq t \leq 300 \\ \sigma_t \varepsilon_t, \sigma_t^2 = 10 + 0.6X_{t-1}^2, \varepsilon_t \sim N(0, 1) & \text{if } 301 \leq t \leq 450 \end{cases} \quad (3.2)$$

$$X_t = \begin{cases} 0.1 + X_{t-1} + \varepsilon_t - 0.8\varepsilon_{t-1}, \varepsilon_t \sim N(0, 1) & \text{if } 1 \leq t \leq 180 \\ 0.1 + X_{t-1} + \varepsilon_t - 0.1\varepsilon_{t-1} - 0.5\varepsilon_{t-2}, \varepsilon_t \sim N(0, 1) & \text{if } 181 \leq t \leq 200 \\ 0.1 + X_{t-1} + \varepsilon_t - 0.8\varepsilon_{t-1}, \varepsilon_t \sim N(0, 1) & \text{if } 201 \leq t \leq 300 \\ 0.1 + X_{t-1} + \varepsilon_t - 0.2\varepsilon_{t-1} - 0.7\varepsilon_{t-2}, \varepsilon_t \sim N(0, 1) & \text{if } 301 \leq t \leq 320 \\ 0.1 + X_{t-1} + \varepsilon_t - 0.8\varepsilon_{t-1}, \varepsilon_t \sim N(0, 1) & \text{if } 321 \leq t \leq 450 \end{cases} \quad (3.3)$$

From the above figure, we can find that the data lead certain types of nonlinear process. In model (3.1) and (3.2), it fluctuates along a constant mean, while the data in model (3.3) illustrates an upward trend.

#### 3.1. Genetic Modeling Analysis

Firstly, we equally divide these time series into 11 generations. In order to emphasize the heredity concept, we let each generation has 10 point overlapped. That is, the last 10 data at the  $i^{\text{th}}$  generation is the same as the first 10 data of the  $(i + 1)$  generation. According to the techniques we discuss above, we find the following result.

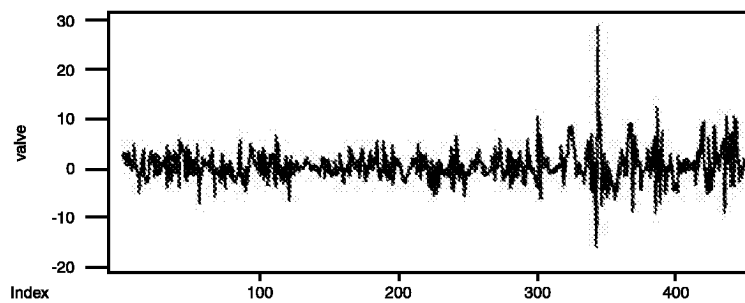


Figure 3.2. Trend for the time series model (3.2) (mean = 0.41, SD = 3.72).



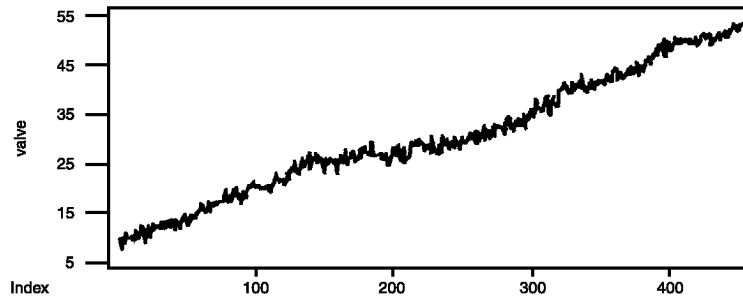


Figure 3.3. Trend for the time series model (3.3) (mean = 30.42, SD = 12.09).

For the model (3.1), we choose AR(1), AR(2), MA(1), MA(2), ARMA(1,1) as our gene models. According to the fitted AIC value, we calculate each membership for candidate models as well as the  $n_i(j)$  values with the significance level  $\alpha = 0.1$  which were shown at Table 3.1. Under the confidence level  $\lambda = 0.8$ , since  $S(1) = 0.82 > 0.8$ , we say that AR(1) is the leading model for the time series (3.1).

The leading model AR (1) has the  $n_i(j)$  value 0 at the generation 4 and 8; it shows that the leading model at this generation is not the best-fitted model. According to the identification and decision algorithm, the 4<sup>th</sup> and the 8<sup>th</sup> generation, which are 120 ~ 170 and 180 ~ 330, are not changing periods. This coincides with the original model (3.1).

By equation (2.2), since  $\sum_{j=1}^5 S(j) = 0.82 + 0.18 + 0.18 + 0.18 + 0.27 = 1.63$ , and  $w_1 = \frac{0.82}{1.63} = 0.5031$ ,  $w_2 = \frac{0.18}{1.63} = 0.1104$ ,  $w_3 = \frac{0.18}{1.63} = 0.1104$ ,  $w_4 = \frac{0.18}{1.63} = 0.1104$ ,  $w_5 = \frac{0.27}{1.63} = 0.1656$ , the optimal forecasting  $F_t$  is

$$F_t = 0.5031f_{1,t} + 0.1104f_{2,t} + 0.1104f_{3,t} + 0.1104f_{4,t} + 0.1656f_{5,t}.$$

For the model (3.2), we calculate memberships for each gene model as well as the  $n_i(j)$  value with the significance level  $\alpha = 0.1$  which were shown at Table 3.2. Under the

Table 3.1. Memberships and values  $n_i(j)$  for model (3.1).

	AR(1)	AR(2)	MA(1)	MA(2)	ARMA(1,1)
Generation 1	1.00 (1)	0.89 (0)	0.68 (0)	0.70 (0)	0.76 (0)
Generation 2	1.00 (1)	0.76 (0)	0.76 (0)	0.745 (0)	0.75 (0)
Generation 3	1.00 (1)	0.95 (1)	0.51 (0)	0.63 (0)	0.92 (1)
Generation 4	0.64 (0)	0.79 (0)	1.00 (1)	0.83 (0)	0.78 (0)
Generation 5	1.00 (1)	0.82 (0)	0.57 (0)	0.82 (0)	0.82 (0)
Generation 6	1.00 (1)	0.81 (0)	0.70 (0)	0.76 (0)	0.77 (0)
Generation 7	1.00 (1)	0.89 (0)	0.53 (0)	0.74 (0)	0.87 (0)
Generation 8	0.46 (0)	0.81 (0)	0.90 (1)	0.95 (1)	0.89 (1)
Generation 9	1.00 (1)	0.95 (1)	0.43 (0)	0.69 (0)	0.96 (1)
Generation 10	0.91 (1)	0.80 (0)	0.46 (0)	1.00 (1)	0.81 (0)
Generation 11	1.00 (1)	0.86 (0)	0.50 (0)	0.80 (0)	0.86 (0)
S(j)	0.82	0.18	0.18	0.18	0.27

Table 3.2. Memberships and  $n_i(j)$  values for model (3.2).

	AR(1)	AR(2)	MA(1)	MA(2)	ARIMA(1,1)
Generation 1	0.91 (1)	0.89 (0)	0.91 (1)	0.75 (0)	1.00 (1)
Generation 2	1.00 (1)	0.79 (0)	0.98 (1)	0.74 (0)	0.74 (0)
Generation 3	0.88 (0)	0.91 (1)	0.90 (1)	0.82 (0)	1.00 (1)
Generation 4	0.98 (1)	0.82 (0)	1.00 (1)	0.84 (0)	0.79 (0)
Generation 5	0.95 (1)	0.69 (0)	1.00 (1)	0.75 (0)	0.73 (0)
Generation 6	0.73 (0)	0.96 (1)	0.69 (0)	0.82 (0)	1.00 (1)
Generation 7	1.00 (1)	0.75 (0)	0.92 (1)	0.72 (0)	0.84 (0)
Generation 8	0.76 (0)	0.93 (1)	0.66 (0)	1.00 (1)	0.72 (0)
Generation 9	0.74 (0)	0.98 (1)	0.78 (0)	1.00 (1)	0.88 (1)
Generation 10	0.69 (0)	1.00 (1)	0.75 (0)	0.89 (0)	0.80 (0)
Generation 11	0.82 (0)	0.81 (0)	0.85 (0)	0.80 (0)	1.00 (1)
$S(j)$	0.45	0.45	0.55	0.18	0.45

confidence level  $\lambda = 0.8$ , since  $S(2)$  and  $S(4)$  both equal to 0.55, which is less than the confidence level  $\lambda = 0.8$ . We say that there does not exist a leading model for the time series (3.2). The time series exhibits a very unstable process. In fact, the underlying data came from an ARCH(1) process, our conclusion is very reasonable.

By equation (2.2), since  $\sum_{j=1}^5 S(j) = 0.45 + 0.45 + 0.55 + 0.18 + 0.45 = 2.08$ , and  $w_1 = \frac{0.45}{2.08} = 0.2163$ ,  $w_2 = \frac{0.45}{2.08} = 0.2163$ ,  $w_3 = \frac{0.55}{2.08} = 0.2644$ ,  $w_4 = \frac{0.18}{2.08} = 0.0865$ ,  $w_5 = \frac{0.45}{2.08} = 0.2163$ , the optimal forecasting  $F_t$  is

$$F_t = 0.2163f_{1,t} + 0.2163f_{2,t} + 0.2644f_{3,t} + 0.0865f_{4,t} + 0.2163f_{5,t}.$$

For the model (3.3), we differentiate the underlying time series with first order, and then separate it into 11 generations. Since AR(1) model does not fit well in the model construction, we choose AR(2), MA(1), MA(2), ARIMA(1,1) as our gene models. Table 3.3 shows the membership for candidate models as well as the  $n_i(j)$  values with the significance level  $\alpha = 0.1$ .

Under the confidence level  $\lambda = 0.8$ , since  $S(1) = 0.82 > 0.8$ , we say that MA(1) is the leading model for the time series (3.3).

By equation (2.2), since  $\sum_{j=1}^4 S(j) = 0 + 0.82 + 0.45 + 0.55 = 1.82$ , and  $w_1 = 0$ ,  $w_2 = \frac{0.82}{1.82} = 0.4505$ ,  $w_3 = \frac{0.45}{1.82} = 0.2473$ ,  $w_4 = \frac{0.55}{1.82} = 0.3022$ , the optimal forecasting  $F_t$  is

$$F_t = 0.4505f_{2,t} + 0.2473f_{3,t} + 0.3022f_{4,t}.$$

#### 4. An Empirical Application for Exchange Rate

In Taiwan, the Central Bank had performed the economic policy of fixed exchange rates before 1979. The adoption of managed floating exchange rates policy from 1979 has made

Table 3.3. Memberships and  $n_i(j)$  values for model (3.3).

	AR(2)	MA(1)	MA(2)	ARIMA(1,1)
Generation 1	0.36 (0)	0.98 (1)	0.82 (0)	1.00 (1)
Generation 2	0.35 (0)	1.00 (1)	0.89 (0)	0.85 (0)
Generation 3	0.76 (0)	1.00 (1)	0.72 (0)	0.69 (0)
Generation 4	0.49 (0)	1.00 (1)	0.81 (0)	0.79 (0)
Generation 5	0.39 (0)	0.78 (0)	0.99 (1)	1.00 (1)
Generation 6	0.28 (0)	0.93 (1)	0.97 (1)	1.00 (1)
Generation 7	0.24 (0)	1.00 (1)	0.93 (1)	0.93 (1)
Generation 8	0.43 (0)	0.70 (0)	0.97 (1)	1.00 (1)
Generation 9	0.32 (0)	1.00 (1)	0.92 (1)	0.93 (1)
Generation 10	0.55 (0)	1.00 (1)	0.80 (0)	0.81 (0)
Generation 11	0.65 (0)	1.00 (1)	0.76 (0)	0.74 (0)
S(j)	0	0.82	0.45	0.55

the exchange rates much more important in the macroeconomic system. Especially, many researchers felt interested in the fields of forecasting the exchange rates. In an open economy, the decision and prediction for the exchange rates are the major factors for the government to constitute leading policies. Hence we try to construct a genetic model to forecast the future values of exchange rates. An available time series of spot exchange rates for US dollars in terms of NT dollars is selected to analyze and compare the performances between ARIMA models and genetic models.

Figure 4.1 is a plot of the monthly exchange rates (NT\$/US\$) from June 1993 to April 1999. This data source comes from the Central Bank of Taiwan. It exhibits several non-regular exchange rates in the 71 time series as well as certain structural change. The non-regularity property of exchange rates makes the analysis of time series very difficult. Figure 4.1 presents a large fluctuation and it seems necessary to detect the change period before constructing a suitable model. Hence, we firstly inspect the tendency of the exchange rates in Figure 4.1.

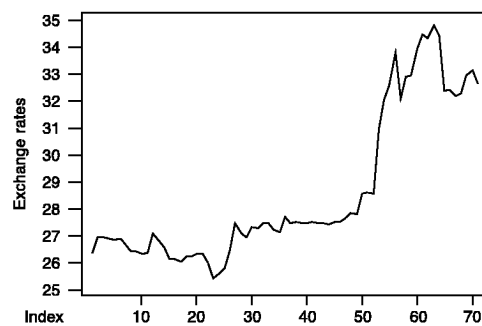


Figure 4.1. Monthly time series for exchange rates from June 1993 to April 1999.

Table 4.1. The membership function for exchange rates.

	AR(2)	MA(1)	MA(2)	ARMA(1,1)
Generation 1	0.90	1.00	0.70	0.54
Generation 2	0.98	0.58	1.00	0.61
Generation 3	0.73	0.71	0.80	1.00
Generation 4	0.72	0.80	0.70	1.00
Generation 5	1.00	0.32	0.69	0.99
Generation 6	0.70	0.70	1.00	0.70

Table 4.2. The  $n_s(j)$  values for exchange rates ( $\alpha = 0.1$ ).

	AR(2)	MA(1)	MA(2)	ARMA(1,1)
Generation 1	1	1	0	0
Generation 2	1	0	1	0
Generation 3	0	0	0	1
Generation 4	0	0	0	1
Generation 5	1	0	0	1
Generation 6	0	0	1	0
S(j)	0.5	0.17	0.33	0.5

In this example, we will use 67 monthly data, from June 1993 to December 1998, in the model construction. The last 4 data are kept for the comparison of forecasting performance.

We separate the time series into 6 generations. In order to emphasize the heredity concept, we let each generation has 5 points overlapping. That is the last 5 data at the  $i^{\text{th}}$  generation is the same as that of  $(i + 1)^{\text{th}}$  generation. We choose AR(2), MA(1), MA(2) and ARMA(1,1) as our gene models, and AIC as our adaptive function. According to the techniques we discuss above and for the significance level  $\alpha = 0.1$ , we illustrate the result in Table 4.1 and Table 4.2.

Under the confidence level  $\lambda = 0.5$ , since  $S(1) = S(4) = 0.50 \geq 0.5$ , we say that AR(2) and ARMA(1, 1) are the leading models for the exchange rates.

#### 4.1. Forecasting Performance

By equation (4.1), since  $\sum_{j=1}^5 S(j) = 0.5 + 0.17 + 0.33 + 0.5 = 1.5$ , and  $w_1 = \frac{0.5}{1.5} = 0.33$ ,  $w_2 = \frac{0.17}{1.5} = 0.12$ ,  $w_3 = \frac{0.33}{1.5} = 0.22$ ,  $w_4 = \frac{0.5}{1.5} = 0.33$ , we have

$$F_t = 0.33f_{1,t} + 0.12f_{2,t} + 0.22f_{3,t} + 0.33f_{4,t}.$$

Table 4.4 illustrates the combined forecasting  $F_t$ , the best fitted ARIMA model for the last generation ( $X_{53} : X_{67}$ ), and the real value. From the table we can find the  $F_t$  has better forecasting performance than the best fitted ARIMA model, MA(2).

Table 4.3. The final results and  $F_t$  for exchange rates.

Model Step	AR(2)	MA(1)	MA(2)	ARMA(1,1)	$F_t$
1	32.36	32.68	33.29	32.35	32.60
2	32.47	33.03	33.21	32.46	32.70
3	32.55	33.03	32.97	32.54	32.69
4	32.61	33.03	32.97	32.59	32.73

Table 4.4. The comparative results for exchange rates.

	Real value	Optimal forecasting $F_t$	Best ARIMA (MA(2))
1999/1	32.30	32.60	33.29
1999/2	32.97	32.69	33.21
1999/3	33.17	32.69	32.97
1999/4	32.70	32.73	32.97

## 5. Conclusion

The weakness of the model-based philosophy of Box-Jenkins modeling clearly resides in the impossibility of satisfying the stationary and the linearity assumptions. However, the uncertainty and fuzziness contained in statistical data is often the obstacle for traditional model construction. If we use quasi-accurate value for cause and effect analysis or quantitative measurement, it will result in biased cause and effect, misleading of decision model, and enlarging difference between prediction and real situation.

This research liberates us from the model-based selection procedure and no assumption of the sample data will be made. Moreover, through the fuzzy identification procedure the genetic modeling process will help us to get an optimal forecasting for a nonlinear time series.

Finally, in spite of the robust forecasting performance for the genetic modeling, there remain some problems for further studies. For example:

- (i) In this research we only use basic ARIMA models as our gene models. In order to get a more extended result, we may include TAR, bilinear or ARCH into our gene model base.
- (ii) The convergence of the algorithm for classification and the proposed statistics have not been well proved, although the algorithms and the proposed statistics are known as fuzzy criteria. This needs further investigation.
- (iii) How to find an optimal test for outliers as well as change periods in the change regime problem?
- (iv) How to find a more comprehensive adaptive function besides AIC?

In order to give the popular questions, such as *duration of a business cycle*, *the occurrence of a turning point*, and *an optimal forecasting*, a satisfactory answer, we

believe the current integrated modeling process with soft computation suggested in this paper will be a worthwhile approach and will stimulate more future empirical work in the nonlinear time series analysis.

### Note

1. Corresponding author: Berlin Wu +886-2-29387387.

### References

- Brown, R., J. Dubin, and J. Evans. (1975). "Techniques for Testing the Constancy of Regression Relationships Over Time," *Journal of the Royal Statistical Society, Ser. B*, 37, 149–163.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.
- Inclan, C. and G. C. Tiao. (1994). "Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance," *Journal of the American Statistical Association* 89(427), 913–924.
- Nguyen, H. and M. Sugeno. (1998). *Fuzzy Modeling and Control*. CRC Press.
- Rukhin, A. (1997). "Change-Point Estimation under Asymmetric Loss," *Statistics & Decisions* 15, 141–163.
- Saatri, T., B. Flores, and J. Valdes. (1989). "Detecting Points of Change in Time Series," *Computers Open Res* 16, 271–293.
- Wu, B. and M. Chen. (1999). "Use Fuzzy Statistical Methods in Change Periods Detection," *Applied Mathematics and Computation* 99, 241–254.
- Wu, B. and C. Chung. (2002). "Using Genetic Algorithms to Parameters (d, r) Estimation for Threshold Autoregressive Models," *Computational Statistics and Data Analysis* 38, 315–330.