Journal of Sound and Vibration (1997) **205**(3), 309–321

ESTIMATION OF THE EXPONENTIAL AUTOREGRESSIVE TIME SERIES MODEL BY USING THE GENETIC ALGORITHM

Z. Shi and H. Aoyama

Division of Intelligent Instrument and Information Integration, Department of Mechanical Engineering, Shizuoka University, 3-5-1 *Johoku*, *Hamamatsu*, *Shizuoka* 432, *Japan*

(*Received* 11 *June* 1996, *and in final form* 17 *March* 1997)

Exponential autoregression (EAR) is a kind of useful non-linear time series model that has properties similar to those of non-linear random vibrations. This model is of autoregressive form with amplitude-dependent coefficients, so parameter estimation is a non-linear optimization problem. To achieve this difficult but important task, this paper introduces a new procedure of the genetic algorithm hybridized with the least squares method to estimate the model. The simulations of both artificial time series and actual data are given to show the efficiency of the proposed approach.

1. INTRODUCTION

An exponential autoregressive time series model, which may give a good fitting of many stochastic processes observed in practical social, economic and engineering systems, was introduced first by Ozaki and Oda [1]. It has been well known as a useful non-linear time series model [2] because some non-linear vibration phenomena such as amplitude-dependent frequency, jump phenomenon and limit cycle occur for this type of model too [3]. In accordance with the general equation of non-linear random vibration:

$$\ddot{x} + f(\dot{x}) + g(x) = \eta, \tag{1}$$

the model was originally of a second order autoregressive form by making the coefficients amplitude-dependent:

$$x_t = (\varphi_1 + \pi_1\, e^{-\gamma x_{t-1}^2})x_{t-1} + (\varphi_2 + \pi_2\, e^{-\gamma x_{t-1}^2})x_{t-2} + e_t, \tag{2}$$

where φ_1, φ_2, π_1, π_2, γ are constants, $\{e_t, t=1, \ldots, N\}$ is a white noise series, and $\{x_t, t = 1, \ldots, N\}$ are the observations under study with a zero mean. It has been shown that this model may exhibit limit cycle behavior under some conditions, as for Van der Pol's equation by ignoring white noise input [4].

The above second order model may be readily extended to a general order model. Thus, a pth order exponential autoregressive model is given as

$$x_t = (\varphi_1 + \pi_1\, e^{-\gamma x_{t-1}^2})\mathrm{x}_{t-1} + \cdots + (\varphi_p + \pi_p\, e^{-\gamma x_{t-1}^2})x_{t-p} + e_t. \tag{3}$$

0022–460X/97/330309 + 13 $25.00/0/sv971048

For the above model, Haggan and Ozaki [4] have shown that the necessary conditions of the existence of a limit cycle are

(1) all the roots of $z^p - \varphi_1 z^{p-1} - \cdots - \varphi_p = 0$ lie inside the unit circle.
(2) Some of the roots of $z^p - (\varphi_1 + \pi_1)z^{p-1} - \cdots - (\varphi_p + \pi_p) = 0$ lie outside the unit circle.

A sufficient condition for the existence of a limit cycle is

(3) $(1 - \sum_{i=1}^{p} \varphi_i)/\sum_{i=1}^{p} \pi_i > 1$ or <0.

However, by comparing with simpler structure and broader usage of the exponential autoregressive model, its identification, that is, the estimation of the order p and the $2p + 1$ coefficients $\{\gamma, (\varphi_i, \pi_i, i = 1, \ldots, p)\}$, is difficult because of the non-linear coefficient γ in the exponential terms of the model. Practical estimation of an exponential autoregressive model is a non-linear optimization problem with the objective function:

$$J(\theta) = \sigma_e^2 = \frac{1}{N-p} \sum_{t=p+1}^{N} \left(x_t - \sum_{i=1}^{p} (\varphi_i + \pi_i \, e^{-\gamma x_{t-1}^2}) x_{t-i} \right)^2, \tag{4}$$

where

$$\theta = (\varphi_1, \ldots, \varphi_p, \pi_1, \ldots, \pi_p, \gamma)$$

Unfortunately, it can be proved that this objective function for the non-linear coefficient γ is not convex, so even optimal methods such as non-linear programming are unavailable for obtaining the global optimal values of the coefficients. As an approximate method, Haggan and Ozaki [4] have given the following estimating procedure.

(1) First fix the γ value that may be determined by some preliminary experiments, then, $\varphi_1, \pi_1, \varphi_2, \pi_2, \ldots, \varphi_p, \pi_p$ are estimated by standard least square regression analysis of x_t on x_{t-1}, $x_{t-1} \, e^{-\gamma x_{t-1}^2}, \ldots$, while the order p is selected by minimizing AIC (Akaike information criterion) [5, 6].
(2) The above analysis is repeated by using a range of γ values, and the AIC criterion is used to select the most suitable value of γ. The values of γ selected are such that $e^{-\gamma x_{t-1}^2}$ does not equal 0 or 1 for most values of x_{t-1}.

In fact, the above approach to find the ''best'' model is only a procedure by trial and error. If there is no *a priori* knowledge of the data under study, a large number of experiments have to be carried out although the final model may still be a local optimum.

In this paper, a novel approach for estimating the exponential autoregressive model is provided by using the genetic algorithm (GA) [7–9], which is a class of global optimization procedure distinguished from other optimization techniques by using concepts from population genetics to guide the search. In this approach, the estimation of the non-linear coefficient γ of the exponential autoregressive model is regarded as an optimization problem to which the genetic algorithm is applied, while the other linear coefficients of the model is estimated by the recursive least squares (RLS) method. Moreover, the order p is also selected by minimizing the AIC value.

The remainder of this paper is the description of estimating the exponential autoregressive model by using the genetic algorithm hybridized with the recursive least squares method, and its simulations of artificial and actual data.

2. ESTIMATION PROCEDURE

2.1. THE SIMPLE GENETIC ALGORITHM

The simple genetic algorithm (SGA) works with a set of strings, namely the population, while each string, namely the chromosome (or individual), represents a possible solution to the particular optimization problem. In the simple GA, each chromosome of the population will be a binary string of length L that corresponds to the problem encoding described in the next subsection. Generally, the population of the first generation is randomly assigned, each individual in the population is then evaluated and assigned a fitness. The population then evolves from generation to generation through the application of natural genetic operators: *reproduction*, *crossover*, and *mutation*.

Reproduction: Reproduction is based on the principle of survival of the fittest. In such a case, a fitness function $f(i)$, $(i = 1, \ldots, M)$ that represents the fitness of the ith chromosome in the current generation must be assigned to each individual in a generation where high values mean good fitting. In this operation, the chromosomes will be selected to survive, and replace the eliminated ones based on the normalized fitness $F(i) = f(i)/\overline{f}$, where $\overline{f}$ is the average fitness of the chromosomes in the current generation:

$$\overline{f} = 1/M \sum_{i=1}^{M} f(i)$$

and M is the number of the chromosomes in each generation.

Crossover: In the genetic algorithm, reproduction directs only the search toward the best fitting individuals but does not create any new individuals. So the crossover operation is introduced to create new individuals. Here the 1-point crossover is considered, in which two new individuals will be recombined through crossing a pair of strings, namely parents, with a certain probability p_c called crossover rate. The crossover rate controls the frequency with which the crossover operation is applied. For example, with the following two strings as parents:

parents 1   0 0 0 0 0 0 0 0   parents 2   1 1 1 1 1 1 1 1

then, two children will be produced in the 1-point crossover operation sited at 4 for instance,

children 1   0 0 0 0 1 1 1 1   children 2   1 1 1 1 0 0 0 0

Mutation: After the crossover, mutation is introduced to insure against premature convergence that would perhaps be a locally optimal solution. This operation simply alters the gene in a chromosome from 1 to 0 or from 0 to 1 with a low probability p_m.

After the operations described above have been carried out, the individuals in a new generation are formed. Figure 1 shows the simple GA's operations from a generation to the next one.

2.2. ENCODING OF THE EXPONENTIAL AUTOREGRESSIVE MODEL ESTIMATION

Usually there are two main components of the genetic algorithm that are problem dependent: the problem encoding and the evaluation function. It is necessary to make clear the problem under study before applying the GA. To estimate the exponential autoregressive model, one must estimate an optimal order p and a set of parameters $\{\gamma, (\varphi_i, \pi_i, i = 1, \ldots, p)\}$, to minimize the objective function. Obviously the exponential autoregressive model has only one non-linear coefficient γ. Whenever γ is determined, the estimation of this model will be reduced to a linear regression problem that can be easily solved. Therefore, this paper provides a procedure to mutually identify the exponential autoregressive model, in which two steps are carried out alternately for a fixed order p, one of which is to optimize γ values by using the genetic algorithm, and the other is to estimate linear parameters $\{\varphi_i, \pi_i, i = 1, \ldots, p\}$ by the recursive least squares method with the given γ values through GA search.

For the problem of encoding the selection of the γ values of the exponential autoregressive model, the unique assumption typically made is that it can be represented by bit strings. This means that the γ value can be encoded and discretized in the range of discretization corresponding to some power of 2. Therefore the γ value represented by a chromosome with L bits will be selected in the real space of $[0, \gamma_{max}]$ with $\gamma_{max} = (2^L - 1)/\delta$, where δ is used to adjust the size of the searching precision and range of the γ value. For example, for a chromosome with length $L = 10$, let $\delta = 102{\cdot}3$, then the γ value will be selected in a range of [0, 10·0] with searching precision 1/102·3. The values of L and δ are determined according to the problem under study.

Besides the coding task, the evaluation function also takes an important role in developing a good simulation in the genetic algorithm. In this procedure, the AIC criterion is used to select the optimal model. As is well known, the smaller the AIC value of the estimated model, the better the model fitting, and the best fitting model may be the one with the minimum AIC value. In practice, the objective function $J(\theta)$ given by equation (4) is employed to be an evaluation function instead of the AIC in estimating the γ value, since the minimization of the AIC is reduced to the minimization of $J(\theta)$ for the fixed order p. On the other hand, the genetic algorithm operates on the principle of natural selection, so each chromosome must be given a fitness that transforms the measure of performance into an allocation of reproductive opportunities directly, the individuals with higher fitness will get a greater chance to survive and produce offspring. Therefore, the fitness function is defined as follows:

$$f(i) = \hat{\sigma}_e^2(max) - \hat{\sigma}_e^2(i) \qquad i = 1, 2, \ldots, M, \tag{5}$$

where $\hat{\sigma}_e^2(max)$ is the maximum residual variance or the maximum objective function value in the current generation, and $\hat{\sigma}_e^2(i)$ is the residual variance of the ith individual (model) in the current generation. Thus, the above definition ensures the fitness function $f(i)$ is directed to searching for the optimum γ value that is consistent with minimizing the objective function of model estimation.

2.3. MODEL ESTIMATION VIA GA HYBRIDIZED WITH RLS

Now one is concerned with the achievement of the exponential autoregressive model estimation by using the hybridization of the genetic algorithm and the recursive least

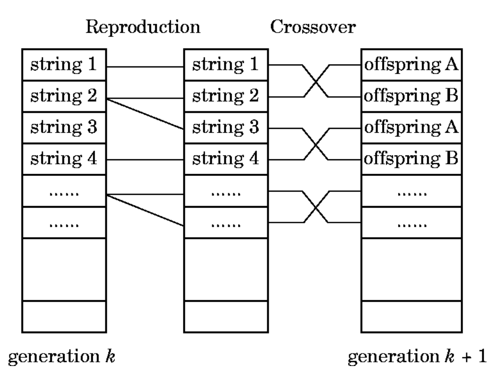

Figure 1. The simple GA operations from a generation to the next one (mutation is not shown).

squares method. As the above description shows, the optimization of the non-linear coefficient γ by the genetic algorithm is not separable from the estimation of other coefficients in the model. Practically, it is helpful to view the estimation of a model's linear coefficients as a necessary step calculating the fitness of each chromosome in the genetic algorithm for understanding this self-organizing modelling procedure. By assuming that the non-linear coefficient of a candidate model decoded from a chromosome in the process of GA search are $\gamma(i)$ for a fixed order p, then this exponential autoregressive model can be described by the following linear regression.

$$x_t = Z^{\mathrm{T}}(t)\theta + e_t, \tag{6}$$

where

$$Z(t) = [x_{t-1}, \ldots, x_{t-p}, \mathrm{e}^{-\gamma(i)x_{t-1}^2}x_{t-1}, \ldots, \mathrm{e}^{-\gamma(i)x_{t-1}^2}x_{t-p}]^{\mathrm{T}}$$

$$\theta = [\varphi_1, \ldots, \varphi_p, \pi_1, \ldots, \pi_p]^{\mathrm{T}}$$

For estimating linear form models of this kind, the following least squares method is used:

$$\hat{\theta}(t+1) = \hat{\theta}(t) + K(t+1)\varepsilon(t+1),$$

$$K(t+1) = P(t)Z(t+1)(1 + Z^{\mathrm{T}}(t+1)P(t)Z(t+1))^{-1},$$

$$P(t+1) = P(t)(1 - K(t+1)Z^{\mathrm{T}}(t+1)), \qquad \varepsilon(t+1) = x_{t+1} - Z^{\mathrm{T}}(t+1)\hat{\theta}(t). \tag{7}$$

With $\hat{\theta}(N)$ as the final estimated coefficients, the model is measured by the evaluation function for its fitting performance which is used to guide the further searching direction of the γ value.

In summary, for the estimation of the exponential autoregressive model with a fixed order p, the proposed procedure consists of the steps:

(1) Initialize all of the tuning parameters of the genetic algorithm: M, L, δ, p_c, p_m and the maximum generation for evolution G_MAX. Also define the initial values $P(0)$ and $\hat{\theta}(0)$ of the recursive least squares method. Encode the initial M chromosomes randomly, set $k = 1$.

(2) Decode the chromosomes into the non-linear coefficients $\gamma(i)$ of the models. Estimate each model by the least squares method, calculate the evaluation function $\hat{\sigma}_e^2(i)$, in turn the fitness and the normalized fitness.

(3) *Reproduction and offspring protection*: If the fitness of the best chromosome in the current generation is less than that of the last generation, the best chromosome of the last generation replaces the worst one in the current generation. Then select the chromosomes of higher fitness to survive and replace the ones of less fitness.

(4) *Crossover*: Each two neighboring chromosomes swap the fragments in a random site determined with the crossover rate p_c.

(5) *Mutation*: Each bit of each chromosome in the current generation mutates with the mutation rate p_m.

(6) $k = k + 1$ if $k > G\_MAX$, then terminate the algorithm, otherwise, go to step (2).

(7) Output all the coefficients and the residual variance of the optimum model of the fixed order p.

Moreover, this estimating procedure is carried out for models of different order p, and these optimum models of different orders are checked by their AIC values calculated by the following definitions [5, 6]:

$$AIC = (N - p)\,(\ln 2\pi + 1 + \ln \hat{\sigma}_e^2) + 2(2p + 1) \tag{8}$$

If $N \gg p$, then the simplification of the above equation leads to

$$AIC = (N - p) \ln \hat{\sigma}_e^2 + 2(2p + 1) \tag{9}$$

Generally, the final model will be the model of the minimum AIC value and simultaneously shares the main properties of the raw data under study. In this paper, the limit cycle behavior of the models of different order p is examined by the model's unforced response using the first p observations of the series under study as the initial values. Furthermore, a comparison of the predicted performance of these candidate models is also used to help select the final model.

3. EXAMPLES

In this section, four examples are presented to illustrate the performance of this new procedure for exponential autoregressive model estimation. The first two examples originate in the existing literature [4], which is convenient for the purpose of comparison. The third example, which is an artificial series generated from a linear autoregressive (AR) model, is used to show the adaptive estimating ability of this procedure, and the last example is an application of the exponential autogressive model to predict the fluctuations in a far-infrared laser. In all these examples the tuning parameters of the genetic algorithm and the initial values of the recursive least squares method are determined as follows: $M = 20$, $L = 10$, $p_c = 0{\cdot}8$, $p_m = 0{\cdot}01$, $\delta = 102{\cdot}3$, $G\_MAX = 1000$, $P(0) = 10^5 I$, $\hat{\theta}(0) = 0{\cdot}0$. Note that a smaller δ value such as 10·23 or 1·023 may be selected instead of 102·3, requiring the estimating procedure to be repeated for a larger search range of γ, when the estimated γ value of the optimum model reaches the maximum value γ_{max} of the search space. The AIC value is calculated using equation (9), except for Example 2 which uses equation (8).

Example 1: 1000 observations are generated from a second order exponential autoregressive model that is known to have the limit cycle behavior:

$$x_t = (1{\cdot}95 + 0{\cdot}23\, e^{-x_{t-1}^2})x_{t-1} - (0{\cdot}96 + 0{\cdot}24\, e^{-x_{t-1}^2})x_{t-2} + e_t,$$

where $\{e_t\}$ is the white noise input with variance 0·001 and mean zero. For this series, 100 replications using the presented procedure for the models of order 1–4 agree with the results shown in Table 1 (where Y means that the model has a limit cycle and N means that it has no limit cycle). Obviously the final model is a second order EAR model with the linear coefficients:

$$\hat{\varphi}_1 = 1{\cdot}949990, \quad \hat{\varphi}_2 = -0{\cdot}959991, \quad \hat{\pi}_1 = 0{\cdot}229747, \quad \hat{\pi}_2 = -0{\cdot}239553,$$

Thus good performance of the procedure has been demonstrated since the estimated model is very similar to the original model in this simulation of artificial data.

TABLE 1

The comparison of the optimum models of different orders for example 1

p	$\hat{\gamma}$	$\hat{\sigma}_e^2$	AIC	Limit cycle
1	162·268	$1{\cdot}824 \times 10^{-1}$	−1693·8	N
2	0·997067	$2{\cdot}296 \times 10^{-6}$	−12948·5	Y
3	0·997067	$2{\cdot}288 \times 10^{-6}$	−12934·7	Y
4	0·987292	$2{\cdot}267 \times 10^{-6}$	−12927·0	Y

TABLE 2

The comparison of the optimum models of different orders for example 2

p	$\hat{\gamma}$	$\hat{\sigma}_e^2$	AIC	Limit cycle
1	35·3861	0·1100	77·2734	N
2	1·3490	0·0465	−15·9017	N
3	1·5542	0·0457	−13·4696	N
4	2·1799	0·0442	−12·9219	N
5	1·7009	0·0440	−9·2491	Y
6	1·6813	0·0438	−5·3735	Y
7	2·2483	0·0410	−8·2240	Y
8	2·1896	0·0376	−13·0660	N
9	2·3167	0·0371	−9·9640	N
10	3·3822	0·0361	−8·2612	N
11	3·8905	0·0321	−16·0562	Y
12	0·2542	0·0289	−22·1464	N
13	0·3619	0·0289	−17·2224	N
14	0·3421	0·0289	−12·6625	N

Example 2: The Canadian lynx data is a well-known time series that has been studied with much interest. It consists of 114 observations describing the annual number of lynx trapped in the Mackenzie river district of Canada over the years 1821 to 1934 (given in [2]). The data is non-linear and its main feature is the strong cyclical behavior. Therefore models fitting to this data are expected to have limit cycles similar to the change of the data. For the logarithmically transformed data, Haggan & Ozaki [4] considered it as a van der Pol type perturbed limit cycle process, and fitted it by an exponential autoregressive model of eleventh order. Applying the proposed approach, the authors also obtained identical results using 100 replications (see Table 2), where the AIC values are calculated according to equation (8) because the lynx data is a small sample time series. In Table 2, the models of order 12 and 13 have no limit cycle behavior, so they are not suitable as the fitting model of the lynx data although they are of smaller AIC values. The reason may be due to the fact that there is a probability of overfitting a small sample in using the AIC criterion. Therefore the final model is obviously the 11th order exponential autoregressive model with $\hat{\gamma} = 3{\cdot}8965$, and the linear coefficients shown in Table 3, which agree well with the results of Haggan & Ozaki [4]. Moreover, this final model also satisfies the conditions for existence of a limit cycle, and its unforced response is shown in Figure 2. It can be seen from Figure 2 that the period of the limit cycle is about 9·5 years which is very similar to that of the lynx data. The comparison of the lynx data and the simulation data using the final model given in Figure 3 also shows that the model fits the lynx data well. The estimated residual variance is 0·032, which is the same as that given by Haggan & Ozaki [4], and also compares favorably with the values found by Campbell and Walker [10], Tong [11] and Tong and Lim [12], 0·039, 0·044 and 0·036, respectively.

TABLE 3

The linear coefficients of the final model of the lynx data

i	1	2	3	4	5	6	7	8	9	10	11
$\hat{\varphi}_i$	1·092	−0·277	0·265	−0·442	0·405	−0·356	0·214	−0·097	0·225	0·065	−0·380
$\hat{\pi}_i$	0·015	−0·487	−0·061	0·290	−0·531	0·599	−0·532	0·302	−0·182	0·180	0·158

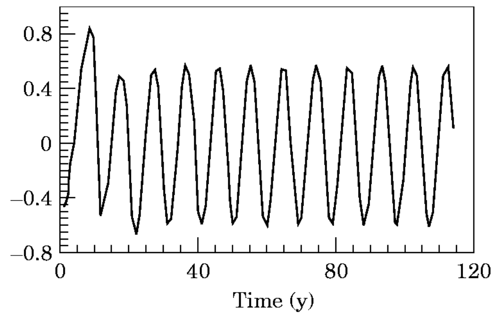

Figure 2. The unforced response of the final model of the lynx data.

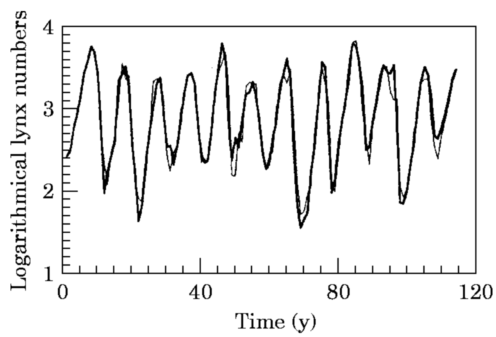

Figure 3. The comparison of the logarithmically transformed data (thick line) and the model's fitting value (thin line).

Example 3: As is well known, the exponential autoregressive model is an extension of the AR model. Whenever the non-linear coefficient γ becomes zero, the EAR model is reduced to a linear AR model. Through taking account of this point, the authors think that an adaptive estimating algorithm for the exponential autoregressive model must be available for both non-linear and linear time series. That is, the non-linear coefficient γ may be estimated to be zero automatically if the data under study is linear, in this case, the final model will be an AR model of the coefficients $\{\varphi_i + \pi_i\}$. To check the ability of the proposed procedure, a series of 1000 data from the following arbitrary second order AR model is generated:

$$x_t = 1{\cdot}1x_{t-1} - 0{\cdot}8x_{t-2} + e_t$$

where $\{e_t\}$ is the white noise with mean zero and variance 0·01. Consequently, 100 replications of the data by the present modelling procedure produce identical results as

TABLE 4
The comparison of the optimum models of different orders for example 3.

p	$\hat{\gamma}$	$\hat{\sigma}_c^2$	AIC	Limit cycle
1	961·877	$3{\cdot}1342 \times 10^{-4}$	−8053·89	N
2	0·000000	$1{\cdot}0527 \times 10^{-4}$	−9130·62	N
3	62·4633	$1{\cdot}0526 \times 10^{-4}$	−9117·62	N
4	59·9218	$1{\cdot}0512 \times 10^{-4}$	−9105·78	N

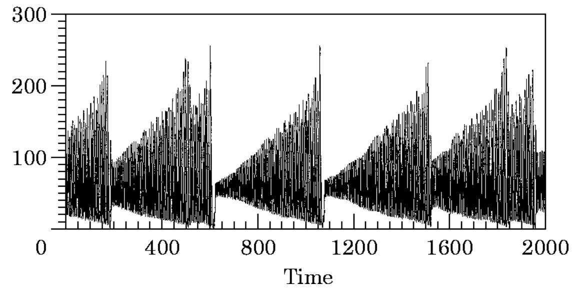

Figure 4. A record of the fluctuations in a far-infrared laser.

shown in Table 4, where the final model is obviously second order with the following estimated coefficients:

$$\hat{\gamma} = 0{\cdot}000000, \quad \hat{\varphi}_1 = 0{\cdot}557373, \quad \hat{\varphi}_2 = -0{\cdot}407774, \quad \hat{\pi}_1 = 0.557373, \quad \hat{\pi}_2 = -0{\cdot}407774.$$

This final model is confirmed as a second order AR model by its similarity to the original model:

$$x_t = 1{\cdot}114746 x_{t-1} - 0{\cdot}815548 x_{t-2} + \varepsilon_t,$$

where $\{\varepsilon_t\}$ is the residual with the variance 0·000105.

Example 4: By the estimating procedure presented, an exponential autoregressive model has been applied to predict the fluctuations in a far-infrared laser as depicted in Figure 4 (The data are available on WWW at: http://www.cs.colorado.edu/~addreas/Time-Series/SantaFe, in files A.dat(the first 1000 points) and A.cont(contains continuation for file A.dat; the first 1000 points of the continuation is used for the prediction test)). This series of 2000 points is obviously changing in a cyclical way similar to non-linear random vibration, so it may be appropriate to use the exponential autoregressive model to fit and in turn predict the series. In this experiment, the modelling procedure presented was applied to estimate the optimum model of the first 1000 data of the series, and then the trained model of the first 1000 points was used to predict one-step-ahead the remaining

TABLE 5

The comparison of the optimum models of different orders for example 4

p	$\hat{\gamma}$	$\hat{\sigma}_e^2$	AIC	Limit cycle	predicting error variance
1	21·016618	1574·71	7360·46	N	163203·20
2	22·580645	866·14	6760·52	N	42525·37
3	0·009775	781·83	6655·65	Y	860·78
4	0·009775	522·22	6251·05	Y	564·19
5	0·009775	462·45	6127·85	Y	517·28
6	0·009775	455·10	6109·80	Y	508·07
7	0·009775	374·47	5914·04	Y	449·07
8	0·009775	352·86	5853·13	Y	424·00
9	0·009775	347·90	5837·24	Y	421·92
10	0·009775	337·88	5806·45	Y	416·68
11	0·009775	335·31	5797·09	Y	413·03
12	0·009775	328·89	5776·18	Y	406·26
13	0·009775	324·37	5760·73	Y	403·57
14	0·009775	315·02	5730·10	Y	385·92
15	0·009775	309·65	5711·40	Y	396·42
16	0·009775	299·55	5677·05	N	811·29

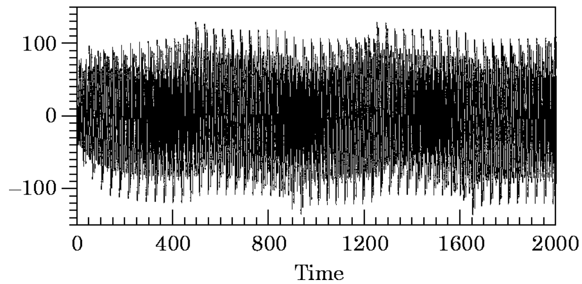

Figure 5. The unforced response of the final model of the series of the fluctuations in a far-infrared laser.

1000 points. Consequently, the optimum models of different orders estimated by the present approach for the first 1000 data (mean $\bar{x} = 59{\cdot}89$ eliminated), and their predicting error variances for the remaining 1000 data, are obtained and shown in Table 5. On the balance of the AIC value, predicting error variance and the non-linear behavior such as limit cycle, the 14th order EAR model of the coefficients given in Table 6 is selected as the final model because of its minimum predicting error variance. The unforced response of the final model is shown in Figure 5, the simulating data ($t = 1, \ldots, 1000$) and the one-step-ahead prediction data ($t = 1001, \ldots, 2000$) of the model are shown in Figure 6, which are very similar to the original data although there are some negative simulating and predicting values in some points of small values that can be modified to zero in practical application.

4. DISCUSSION

The possibility of searching not only for the non-linear coefficient γ but also the exponential autoregressive model order p simultaneously by using the genetic algorithm has been discussed.

At first, for the problem of encoding the selecting of γ and p of the exponential autoregressive model simultaneously, one assumes that γ and p are represented successively from left to right by the total L bits genes of a chromosome, where the γ value is represented by the first L_1 bits, and p is described by the next L_2 bits, with $L = L_1 + L_2$. According to this kind of encoding, order p will be searched in the integer space of $[0, p_{max}]$ with $p_{max} = 2^{L_2} - 1$, and γ will be selected in the real space of $[0, \gamma_{max}]$ with γ_{max} defined as $\gamma_{max} = (2^{L_1} - 1)/\delta$. Hence there is no problem in coding when using the genetic algorithm.

On the other hand, by using the AIC criterion as the evaluation function, the fitness function can be defined as:

$$f(i) = AIC(max) - AIC(i), \qquad i = 1, \ldots, M, \tag{10}$$

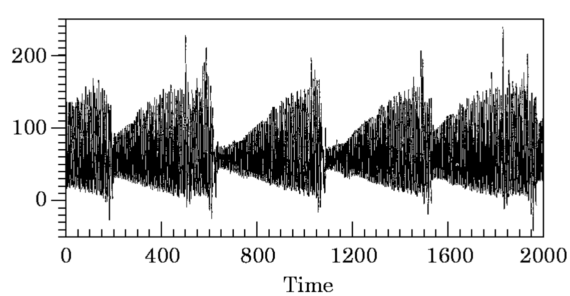

Figure 6. The simulating data (the first 1000 points) and the predicting data (the second 1000 points) of the final model of the fluctuations in a far-infrared laser.

TABLE 6

The linear coefficients of the final model of the first 1000 *points of the fluctuations in a far-infrared laser*

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\hat{\varphi}_i$	0·533	−0·643	0·116	−0·264	−0·055	−0·184	0·236	0·380	−0·145	0·366	−0·065	0·202	0·051	0·180
$\hat{\pi}_i$	−1·231	0·265	0·292	−1·512	−0·089	1·472	−0·748	0·758	−0·100	−0·402	0·110	0·233	−0·860	−0·105

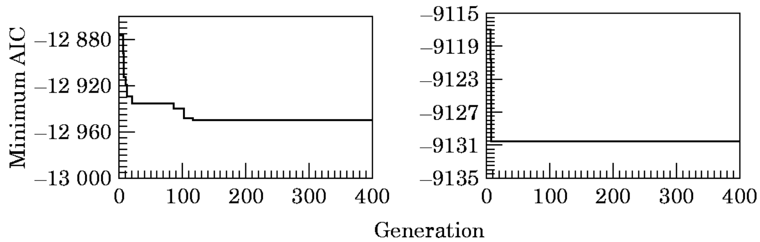

Figure 7. The minimum AIC value in each generation for an arbitrary estimation. (a) Example 1; (b) example 3.

where

$$AIC(i) = (N - p(i))(\ln 2\pi + 1 + \ln \hat{\sigma}_e^2(i)) + 2(2p(i) + 1), \qquad \hat{\sigma}_e^2(i) = \frac{1}{N - p(i)} \sum_{t = p(i)+1}^{N} \hat{e}_t^2(i). \tag{11}$$

AIC(max) is the maximum AIC value in the current generation of M individuals, $p(i)$, $\hat{e}_t(i)$ and $\hat{\sigma}_e^2(i)$ are the order, residual of time t and residual variance of the ith individual (model) in the current generation, respectively.

With the above definition, the procedure presented of the genetic algorithm hybridized with the recursive least squares method may also be used to search for the "optimum" model of the minimum AIC value. With the same tuning parameter of the genetic algorithm and initial values of the recursive least squares method as those in the third section, except for $L = 14$, $L_1 = 10$, and $L_2 = 4$, 100 replications for the data of Example 1 and Example 3 were carried out, and the procedure searched optimum models identical to their final models selected in the third section of this paper. Figure 7(a) and Figure 7(b) show the minimum AIC value in each generation for an arbitrary experiment of Example 1 and Example 3, respectively, for proving the convergence of the proposal. Thus, the possibility of estimating γ and p of the exponential autoregressive model in self-organization by the genetic algorithm is verified. However, sometimes the "optimum" model of minimum AIC value may not be always the best fitting model especially in non-linear time series modelling. It has been argued by Tong [13] that in determining the final choice of model one need not adhere strictly to the values of the "structural parameters" selected by the AIC criterion, rather the AIC criterion is used as guide to select a relatively small subclass of plausible models that may then be examined for certain special properties (such as limit cycle behavior). This can be also seen from the results of numerical examples 2 and 4. It is indeed a fact that there is no criterion available for correctly selecting the final time series model, although the AIC criterion has been very widely accepted for model order selection. Therefore, to estimate the exponential autoregressive model in complete self-organization by using the genetic algorithm, it may be necessary to add some other check of the candidate model properties besides the performance measurement in terms of AIC criterion. This makes the problem more complex, and will be discussed in other presentations.

5. CONCLUSION

A novel method used for optimal estimation of the exponential autoregressive model has been presented. In this approach, the non-linear coefficient γ of the model was searched by using the simple genetic algorithm. Consequently, the estimation of the other coefficients is reduced to a linear regression problem, and is achieved by the recursive least squares method. Compared with trial and error [4] procedure the present approach is self-organizing and globally optimizing with greater accuracy and less calculation.

This approach is simple in principle, but the results are exciting, which can be seen from the examples. It is easy to use and it is not difficult to write an automatic modelling procedure. Moreover, the computational consumption of the procedure depends largely on the computational complexity of the recursive least squares solution and approximately $7(2p)^2 \times N \times M \times G\_MAX$ arithmetic operations are needed for one estimation of the model of order p. However, it is possible to introduce more efficient methods instead of the least squares to estimate the linear coefficients of the model, which may further save CPU time and improve the performance of the approach presented for various non-linear time series.

ACKNOWLEDGMENTS

The authors are grateful to the referees for their valuable comments that improved the presentation of the paper.

REFERENCES

1. T. Ozaki and H. Oda 1978 *Information and System*. Oxford: Pergamon. Non-linear time series model identification by Akaike's Information Criterion, Editor B. Dubission, 83–91.
2. M. B. Priestly 1988 *Non-linear and Non-stationary Time Series Analysis*. London: Academic Press.
3. T. Ozaki 1980 *Journal of Applied Probability* **17**, 84–93. Non-linear time series models for non-linear random vibrations.
4. V. Haggan and T. Ozaki 1981 *Biometrika* **68**, 189–196. Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model.
5. H. Akaike 1974 *IEEE Transactions on Automatic Control* **19**, 716–723. A new look at the statistical model identification.
6. Y. Sakamoto, M. Ishiguro and G. Kitagawa 1986 *Akaike Information Criterion Statistics*. Tokyo: D. Reidel Publishing Company.
7. D. E. Goldberg 1989 *Genetic Algorithms in Search, Optimization, and Machine Learning*. New York: Addison-Wesley.
8. J. Holland 1975 *Adaption in Natural and Artificial Systems*. University of Michigan Press.
9. D. Whitley 1994 *Statistics and Computing* **4**, 65–85. A genetic algorithm tutorial.
10. M. J. Campbell and A. M. Walker 1977 *Journal of the Royal Statistical Society* **39**, 411–431. A survey of statistical work on the Mackenzie river series of annual Canadian lynx trappings for the years 1821–1934 and a new analysis.
11. H. Tong 1977 *Journal of the Royal Statistical Society* **39**, 432–436. Some comments on the Canadian lynx data.
12. H. Tong and K. S. Lim 1980 *Journal of the Royal Statistical Society* **42**, 245–292. Threshold autoregressive, limit cycles and cyclical data.
13. H. Tong 1983 *Threshold Models in Non-linear Time Series Analysis*. New York: Springer-Verlag.