

ESTIMATING THE EMBEDDING DIMENSION AND DELAY TIME OF CHAOTIC TIME SERIES BY AN AUTOREGRESSIVE MODEL

Yonemoto, Koji
Graduate School of Mathematics, Kyushu University

Yanagawa, Takashi
Graduate School of Mathematics, Kyushu University

<https://doi.org/10.5109/13503>

出版情報 : Bulletin of informatics and cybernetics. 33 (1/2), pp.53-62, 2001-12. Research
Association of Statistical Sciences

バージョン :

権利関係 :



ESTIMATING THE EMBEDDING DIMENSION AND DELAY TIME OF CHAOTIC TIME SERIES BY AN AUTOREGRESSIVE MODEL

By

Koji YONEMOTO* and Takashi YANAGAWA†

Abstract

A method of estimating embedding dimension and delay time of chaotic dynamical system with dynamic noise is proposed recently by one of the authors (Fueda and Yanagawa (2001)). It is shown theoretically that the proposed estimators are consistent. However, it is found that the estimators often fail to select the correct embedding dimension and delay time. They are modified in the present paper and shown work well by simulation.

Key Words and Phrases: chaotic time series; cross-validation; dynamic noise; Henon map; Nadaraya-Watson kernel type estimator

1. Introduction

The determination of the embedding dimension and delay time has been a central issues of chaotic dynamical systems. Fraser and Swinney (1986) used the average mutual information to determine the delay time. For the embedding dimension Grassberger and Procaccia (1983) introduced the correlation integral, and Kennel, Brown and Abarbanel (1992) introduced the false nearest neighbors approach. Number of papers have been published that discussed these issues, see, for example Abarbanel (1995). Deterministic systems are exclusively considered in these literature, giving almost no attention to noises that might well exist in an observational data. Cheng and Tong (1994) proposed a method of estimating embedding dimension and delay time from chaotic time series with dynamic noise and the method is recently generalized by Fueda and Yanagawa (2001). They proved the consistency of the proposed estimators. However, we find that the procedure often fails in practice, and that the failure comes from the characteristic of chaos itself, that is, the sensitive dependence of initial conditions. More precisely, as the procedure employs the minimum Cross-Validation (CV) criterion, we find that the minimum is sensitive to initial values. We accomodate the sensitivity in this paper by introducing several rules in the minimum CV criterion. It is shown by simulation that the new procedure works well. In section 2, the Fueda and Yanagawa method is briefly reviewed. New rules are introduced in the minimum CV criterion in section 3 and results of simulation and application are given in section 4 and section 5, respectively. Finally it is concluded that, although the proposed method is developed under the framework of non-linear autoregressive model, it is applicable not only to the data with dynamic noise, but also to those data without dynamic noise, i.e. the deterministic chaotic data.

* Graduate School of Mathematics, Kyushu University, yonemoto@math.kyushu-u.ac.jp

† Graduate School of Mathematics, Kyushu University, yanagawa@math.kyushu-u.ac.jp

2. The model and estimator

2.1. The embedding dimension and delay time

Let $\{X_t\} \subset \mathbb{R}$ be discrete-time stationary time series with $EX_t^2 < \infty$ which is generated by

$$X_t = F(X_{t-\tau}, X_{t-2\tau}, \dots, X_{t-d\tau}) + \varepsilon_t, \quad (1)$$

where ε_t is the dynamic noise, and it is assumed that for any t ,

$$E[\varepsilon_t | \mathcal{A}_1^{t-1}(X)] = 0, \text{ almost surely.} \quad (2)$$

and

$$E[\varepsilon_t^2 | \mathcal{A}_1^{t-1}(X)] = \sigma^2, (\sigma > 0), \text{ almost surely,}$$

where $\mathcal{A}_s^t(X)$ denotes the sigma algebra generated by (X_s, \dots, X_t) . F is assumed to be unknown chaotic nonlinear function. Note that from (1) and (2), it follows that

$$F(X_{t-\tau}, X_{t-2\tau}, \dots, X_{t-d\tau}) = E[X_t | X_{t-\tau}, X_{t-2\tau}, \dots, X_{t-d\tau}]$$

with $E[\varepsilon_t | X_{t-\tau}, X_{t-2\tau}, \dots, X_{t-d\tau}] = 0$.

The time series $\{X_t\}$ is said to have the embedding dimension d_0 with the delay time τ_0 (Yonemoto and Yanagawa, 1998) if and only if there exist non-negative integers $d_0 < \infty$ and $\tau_0 < \infty$ such that

$$E[X_t | X_{t-\tau}, X_{t-2\tau}, \dots, X_{t-d\tau}] \neq E[X_t | X_{t-\tau_0}, X_{t-2\tau_0}, \dots, X_{t-d_0\tau_0}] \text{ a.e.}$$

for any $d < d_0$, and any $\tau > 0$, and

$$E[X_t | X_{t-\tau}, X_{t-2\tau}, \dots, X_{t-d\tau}] = E[X_t | X_{t-\tau_0}, X_{t-2\tau_0}, \dots, X_{t-d_0\tau_0}] \text{ a.e.}$$

for any $(d, \tau) \in \mathbf{B}(d_0, \tau_0)$, where $\mathbf{B}(d_0, \tau_0) = \{(d, \tau) | \{\tau_0, 2\tau_0, \dots, d_0\tau_0\} \subset \{\tau, 2\tau, \dots, d\tau\}\}$.

2.2. Estimation of the embedding dimension and delay time

Let $\{X_1, \dots, X_N\}$ be a set of observed data. For positive integers d, τ and $L \geq d\tau$, put

$$CV(d, \tau) = \frac{1}{N-L+1} \sum_{t=L}^N (X_t - \hat{F}_{t(d, \tau)}(X_{t-\tau}, \dots, X_{t-d\tau}))^2,$$

where $\hat{F}_{t(d, \tau)}$ denotes the Nadaraya - Watson type estimated regression function (Nadaraya, 1964; Watson, 1964) with the t -th point deleted, that is,

$$\hat{F}_{t(d, \tau)}(z) = \frac{1}{N-L} \sum_{s=L, s \neq t}^N X_s K_{d, h}(z - (X_{s-\tau}, \dots, X_{s-d\tau})) (\hat{f}_{t(d, \tau)})^{-1},$$

where the summation over s omit t in each case, and

$$\hat{f}_{t(d, \tau)}(z) = \frac{1}{N-L} \sum_{s=L, s \neq t}^N K_{d, h}(z - (X_{s-\tau}, \dots, X_{s-d\tau})),$$

and

$$K_{d,h}(z) = \frac{1}{h^d} \prod_{i=1}^d \phi\left(\frac{z_i}{h}\right), \quad (z = (z_1, z_2, \dots, z_d))$$

where $\phi(\cdot)$ is taken as a density function of a standard normal distribution in this paper. In estimating the embedding dimension d_0 and delay time τ_0 which are defined in the previous section, Fueda and Yanagawa (2001) introduced the following estimation procedure.

- (F-Y 1) Give sufficiently large integers $D(> d_0)$ and $T(> \tau_0)$, and set $L = DT$. Set $h = h(c) = c \times N^{-1/(2d+1)}$. For each $d \in \{1, 2, \dots, D\}$ and $\tau \in \{1, 2, \dots, T\}$, compute $CV(d, \tau)$.
- (F-Y 2) For each $\tau \in \{1, 2, \dots, T\}$, minimize $CV(d, \tau)$ with respect to $d \in \{1, 2, \dots, D\}$, and denote the minimizer as $\hat{d}_0(\tau)$.
- (F-Y 3) Find $\hat{d}_0 = \min_{\tau} \hat{d}_0(\tau)$ and $\hat{\tau}_0 = \operatorname{argmin}_{\tau} \hat{d}_0(\tau)$.

They proved that \hat{d}_0 and $\hat{\tau}_0$ are consistent estimators of d_0 and τ_0 as $N \rightarrow \infty$. Note that $\hat{\tau}_0$ may not be uniquely determined.

3. Minimization procedure

We simulated the above procedure as follows: For each $d \in \{1, 2, \dots, D\}$ and $\tau \in \{1, 2, \dots, T\}$, first computed $CV(d, \tau)|_{h=h(c)}$ by changing the value c in an appropriate class C , set $\hat{CV}(d, \tau) = \min_{c \in C} CV(d, \tau)|_{h=h(c)}$, and then used this $\hat{CV}(d, \tau)$ as $CV(d, \tau)$ in (F-Y 2) and (F-Y 3). However, it is shown that the estimators \hat{d}_0 and $\hat{\tau}_0$ thus obtained often fail to estimate the true value. See, for example, Table 2 column 2, where no correct estimates of the embedding dimension and delay time are given when $\sigma^2 = 0.01$. It is found that the procedure fails because the minimum CV criterion represented in (F-Y 2) is sensitive to an initial value when F is a chaotic map, and does not provide distinctive minimum value. We propose to modify the procedure as follows.

1. The same as (F-Y 1).
2. For each $\tau \in \{1, 2, \dots, T\}$ and $d \in \{1, 2, \dots, D\}$ minimize $CV(d, \tau)|_{h=h(c)}$ with respect to $c \in C$ by the method illustrated below and put

$$\hat{CV}(d, \tau) = \min_{c \in C} CV(d, \tau)|_{h=h(c)}.$$

3. Then select $d \in \{1, 2, \dots, D\}$ and $\tau \in \{1, 2, \dots, T\}$ which attain the 'minimum' value of $\{\hat{CV}(d, \tau) : \tau \in \{1, 2, \dots, T\}, d \in \{1, 2, \dots, D\}\}$ as estimators of the embedding dimension \hat{d}_0 and delay time $\hat{\tau}_0$ based on the procedure given below.

The detail of the minimization of $CV(d, \tau)|_{h=h(c)}$ with respect to $c \in C$ is given as follows

- a. Give a large real number c_{max} , and for each $d \in \{1, 2, \dots, D\}$ and $\tau \in \{1, 2, \dots, T\}$, compute $CV(d, \tau)|_{h=h(c)}$ and $CV(d, \tau)|_{h=h(c+0.1)}$ starting from $c = 0.1$, compare these values, and if $c < c_{max}$ and $CV(d, \tau)|_{h=h(c)} \geq CV(d, \tau)|_{h=h(c+0.1)}$ then put $c = c + 0.1$ and repeat the computation; else if, stop and decide $c = c_1$.

b. In the neighborhood of c_1 , find $c_2 = \operatorname{argmin} CV(d, \tau)_{h=h(c)}$ in the class $c \in C_1 = \{c_1 - 0.09, c_1 - 0.08, \dots, c_1 + 0.09\}$.

c. Furthermore, in the neighborhood of c_2 , compute

$$\hat{CV}(d, \tau) = \min_c \{CV(d, \tau)_{h=h(c)} : c \in C_2 = \{c_2 - 0.009, \dots, c_2 + 0.009\}\}.$$

The procedure for selecting \hat{d}_0 and $\hat{\tau}_0$ is as follows.

d. Put $CV^*(\tau) = \min\{\hat{CV}(d, \tau) : d \in \{1, 2, \dots, D\}\}$ for each $\tau \in \{1, 2, \dots, T\}$, and find for given $\varepsilon > 0$

$$\hat{d}(\tau) = \min\{d : |\hat{CV}(d, \tau) - CV^*(\tau)| < \varepsilon\}.$$

e. Next put $CV^* = \min\{\hat{CV}(\hat{d}(\tau), \tau) : \tau \in \{1, 2, \dots, T\}\}$, and then find

$$\hat{d}_0 = \min\{\hat{d}(\tau) : |\hat{CV}(\hat{d}(\tau), \tau) - CV^*| < \varepsilon\}$$

and

$$\hat{\tau}_0 = \operatorname{argmin}\{\hat{d}(\tau) : |\hat{CV}(\hat{d}(\tau), \tau) - CV^*| < \varepsilon\}.$$

The smallest $\hat{\tau}_0$ is employed when $\hat{\tau}_0$ is not unique. Step e is introduced to accommodate the sensitivity of CV^* to initial values. The selection of ε is important in the above procedure. We examine the performance of the procedure when ε is $CV^*/10$ and $CV^*/20$ in the simulations.

4. Simulation

The following models are employed in the simulations:

Model 1 : $X_t = 1 - 1.4X_{t-1}^2 + 0.3X_{t-2} + \varepsilon_t$ with $d_0 = 2$ and $\tau_0 = 1$

Model 2 : $X_t = -0.1X_{t-1} + (-1 + 2\exp(-X_{t-1}^2))X_{t-3} + \varepsilon_t$ with $d_0 = 3$ and $\tau_0 = 1$

Model 3 : $X_t = -0.1X_{t-2} + (-1 + 2\exp(-X_{t-2}^2))X_{t-6} + \varepsilon_t$ with $d_0 = 3$ and $\tau_0 = 2$

Model 4 : $X_t = 0.5X_{t-1} + (-0.5 + 2\exp(-X_{t-1}^2))X_{t-7} + \varepsilon_t$ with $d_0 = 7$ and $\tau_0 = 1$

$N + 1000$ points are generated from each model, the first 1000 points are discarded, and then remaining points are used as the data of size N from the model. The details of the simulation and its results are given in the following subsections. The embedding dimension and delay time are estimated based on the data by the proposed method and also by the Fueda and Yanagawa method (F-Y method). The process is repeated 100 times and frequency distributions of the estimated embeddig dimension are obtained. The computation of each trial takes tremendous time even with super computer at the Institute of Statistical Mathematics, therefore 100 is our maximum possible repetition. The results below would not change substantially even if the repetition could be increased.

4.1. Model 1

We consider Model 1 which is called stochastic Henon map. For each $\sigma^2 \in \{0, 0.005, 0.01\}$ and $N \in \{1000, 2000, 4000\}$, data $x_i(t \geq 3)$ are generated by giving initial values (x_2, x_1) which are selected randomly from quadrilateral $ABCD$, where $A = (-1.33, 1.4)$, $B = (1.32, 0.443)$, $C = (1.245, -0.466)$, $D = (-1.06, -1.666)$. This quadrilateral is introduced in Henon (1976) as a region where x_i does not diverge as $t \rightarrow \infty$ when $\sigma^2 = 0$. When $\sigma^2 \neq 0$, divergence could occur even if the initial values are selected from the quadrilateral. To avoid it we discarded x_3, x_4, \dots, x_i if $|x_i - x_{i-1}| > 3$ for some $i \leq 12$, and generated new x_3, x_4, \dots using the same initial values. Also we discarded $x_{i-9}, x_{i-8}, \dots, x_i$ if $|x_i - x_{i-1}| > 3$ for some $i \geq 13$ and generated new x_{i-9}, x_{i-8}, \dots using (x_{i-11}, x_{i-12}) , but keeping $x_1, x_2, \dots, x_{i-10}$. By standardizing the data and setting $D=T=5$, embedding dimension and delay time are estimated. Table 1 summarizes the frequency distributions of \hat{d}_0 and $\hat{\tau}_0$ in 100 trials. The second column of the table shows the results of the F-Y method. The results of the proposed method are given in the third and fourth columns, with $\varepsilon = CV^*/10$ in the third column and $\varepsilon = CV^*/20$ in the fourth column.

N	F-Y method	proposed method	
		$\varepsilon = CV^*/10$	$\varepsilon = CV^*/20$
1000	$(2, 1) \times 100$	$(2, 1) \times 100$	$(2, 1) \times 100$
2000	$(2, 1) \times 100$	$(2, 1) \times 100$	$(2, 1) \times 100$
4000	$(2, 1) \times 100$	$(2, 1) \times 100$	$(2, 1) \times 100$

$$\sigma^2 = 0$$

N	F-Y method	proposed method	
		$\varepsilon = CV^*/10$	$\varepsilon = CV^*/20$
1000	$(1, 5) \times 97$ $(2, 1) \times 3$	$(2, 1) \times 100$	$(2, 1) \times 100$
2000	$(1, 5) \times 79$ $(2, 1) \times 21$	$(2, 1) \times 100$	$(2, 1) \times 100$
4000	$(2, 1) \times 92$ $(1, 5) \times 8$	$(2, 1) \times 100$	$(2, 1) \times 100$

$$\sigma^2 = 0.005$$

N	F-Y method	proposed method	
		$\varepsilon = CV^*/10$	$\varepsilon = CV^*/20$
1000	$(1, 5) \times 100$ $(2, 1) \times 0$	$(2, 1) \times 100$	$(2, 1) \times 100$
2000	$(1, 5) \times 99$ $(2, 1) \times 1$	$(2, 1) \times 100$	$(2, 1) \times 100$
4000	$(1, 5) \times 85$ $(2, 1) \times 15$	$(2, 1) \times 100$	$(2, 1) \times 100$

$$\sigma^2 = 0.01$$

TABLE 1 Frequency distributions of $(\hat{d}_0, \hat{\tau}_0)$ in 100 trials for Model 1, where the bold face represents the correct estimates.

The table shows that the correct estimates are accomplished by both methods when $\sigma^2 = 0$; but the F-Y method becomes poorer as the values of σ^2 increases. For example, it estimates the true value only 15 times in 100 trials when $\sigma^2 = 0.01$ even when

$N = 4000$; whereas the proposed method estimates the true value without errors in all cases.

4.2. Model 2

Next, we consider Model 2. For each $\sigma^2 \in \{0, 0.01\}$ and $N \in \{1000, 2000, 4000\}$, data are generated by giving initial values x_1, x_2, x_3 selected from $(-1, 1)$ randomly. Standardize the data, and by putting $D = 6$ and $T = 5$, the embedding dimension and delay time are estimated. Table 2 summarizes the frequency distributions of \hat{d}_0 and $\hat{\tau}_0$. The second column of the table shows the results of the F-Y method. The results of the proposed method are given in the third column when $\varepsilon = CV^*/10$ and in the fourth column when $\varepsilon = CV^*/20$.

N	F-Y method	proposed method	
		$\varepsilon = CV^*/10$	$\varepsilon = CV^*/20$
1000	(3, 1) \times 46 (3, 2) \times 34 (4, 4) \times 20	(3, 1) \times 65 (4, 4) \times 34 (3, 2) \times 1	(3, 1) \times 43 (4, 4) \times 38 (5, 5) \times 15 (4, 1) \times 3 (3, 2) \times 1
2000	(3, 1) \times 74 (4, 4) \times 21 (3, 2) \times 5	(3, 1) \times 100	(3, 1) \times 100
4000	(3, 1) \times 81 (3, 2) \times 25 (4, 4) \times 24	(3, 1) \times 100	(3, 1) \times 100

$$\sigma^2 = 0$$

N	F-Y method	proposed method	
		$\varepsilon = CV^*/10$	$\varepsilon = CV^*/20$
1000	(1, 3) \times 99 (2, 3) \times 1 (3, 1) \times 0	(1, 3) \times 94 (3, 1) \times 6	(3, 1) \times 51 (1, 3) \times 49
2000	(1, 3) \times 99 (2, 3) \times 1 (3, 1) \times 0	(1, 3) \times 91 (3, 1) \times 9	(3, 1) \times 65 (1, 3) \times 35
4000	(1, 3) \times 99 (2, 3) \times 1 (3, 1) \times 0	(1, 3) \times 87 (3, 1) \times 13	(3, 1) \times 95 (1, 3) \times 5

$$\sigma^2 = 0.01$$

TABLE 2 Frequency distributions of $(\hat{d}_0, \hat{\tau}_0)$ in 100 trials for Model 2, where the bold face represents the correct estimates.

The table shows that when $\sigma^2 = 0$, the F-Y method estimates the true value at most 74 times in 100 trials for $N = 1000, 2000, 4000$; and that when $\sigma^2 = 0.01$, it gives completely wrong estimates. On the other hand, the proposed method estimates the true value without errors when $\sigma^2 = 0$ and N is 2000 and 4000, and that it estimates the true value 95 times in 100 trials when $N = 4000$, $\varepsilon = CV^*/20$, and $\sigma^2 = 0.01$. Note that when $\varepsilon = CV^*/10$, $N = 4000$ and $\sigma^2 = 0.01$ the table shows that only 13 correct estimates in 100 trials by our method, showing that the method is sensitive to the selection of ε .

4.3. Model 3

We consider Model 3. For each $\sigma^2 \in \{0, 0.01\}$ and $N \in \{1000, 2000, 4000\}$, data are generated by giving initial values x_1, x_2, \dots, x_6 selected from $(-1, 1)$ randomly. Standardize the data, and by putting $D = 6$ and $T = 5$, the embedding dimension and delay time are estimated. Table 3 summarizes the frequency distributions of \hat{d}_0 and $\hat{\tau}_0$. The second column of the table shows the results of the F-Y method. The results of the proposed method are given in the third and fourth columns, when $\varepsilon = CV^*/10$ and $\varepsilon = CV^*/20$, respectively.

N	F-Y method	proposed method	
		$\varepsilon = CV^*/10$	$\varepsilon = CV^*/20$
1000	(3, 2) × 51	(3, 2) × 96	(3, 2) × 93
	(3, 4) × 29		
	(6, 4) × 9		
	(5, 3) × 6		
	(4, 2) × 2		
	(6, 2) × 1		
	(5, 4) × 1		
2000	(3, 2) × 36	(3, 2) × 94	(3, 2) × 91
	(3, 4) × 26		
	(6, 4) × 13		
	(5, 3) × 12		
	(5, 5) × 6		
	(4, 2) × 3		
	(5, 2) × 2		
4000	(3, 2) × 36	(3, 2) × 100	(3, 2) × 100
	(3, 4) × 30		
	(5, 3) × 16		
	(4, 3) × 9		
	(6, 4) × 8		
	(4, 2) × 1		

$$\sigma^2 = 0$$

N	F-Y method	proposed method	
		$\varepsilon = CV^*/10$	$\varepsilon = CV^*/20$
1000	(2, 3) × 98	(2, 3) × 68	(3, 2) × 83
	(3, 2) × 2	(3, 2) × 32	(2, 3) × 15
2000	(2, 3) × 100	(2, 3) × 67	(3, 2) × 98
	(3, 2) × 0	(3, 2) × 33	(2, 3) × 2
4000	(2, 3) × 98	(2, 3) × 58	(3, 2) × 100
	(3, 2) × 2	(3, 2) × 42	

$$\sigma^2 = 0.01$$

TABLE 3 Frequency distributions of $(\hat{d}_0, \hat{\tau}_0)$ in 100 trials for Model 3, where the bold face represents the correct estimates.

The table shows that the F-Y method estimates the true value at most 51 times in 100 trials for $N = 1000, 2000, 4000$ when $\sigma^2 = 0$, and that the method gives correct estimates at most 2 times in 100 trials when $N = 1000, 2000, 4000$ and $\sigma^2 = 0.01$; whereas the proposed method estimates at least 91 times in 100 trials for $N = 1000, 2000, 4000$ when $\sigma^2 = 0$; and when $\sigma^2 = 0.01$ it estimates the true value more than 83 times in 100 trials for $N = 1000, 2000, 4000$ if $\varepsilon = CV^*/20$.

4.4. Model 4

Finally, we consider Model 4. For each $\sigma^2 \in \{0, 0.01\}$ and $N \in \{3000, 6000\}$, data are generated by giving initial values x_1, x_2, \dots, x_7 selected from $(-1, 1)$ randomly. The embedding dimension and delay time are estimated by standardize the data and putting $D = 12$ and $T = 5$. Table 4 summarizes the frequency distribution of \hat{d}_0 and $\hat{\tau}_0$ in 100 trials in the same way as previous tables.

N	F-Y method	proposed method	
		$\varepsilon = CV^*/10$	$\varepsilon = CV^*/20$
3000	$(9, 1) \times 63$	$(8, 2) \times 80$ $(8, 1) \times 20$ $(7, 1) \times 0$	$(9, 1) \times 80$ $(8, 1) \times 20$ $(7, 1) \times 0$
	$(9, 2) \times 17$		
	$(6, 3) \times 9$		
	$(7, 3) \times 5$		
	$(7, 2) \times 4$		
	$(5, 3) \times 1$		
	$(5, 5) \times 1$		
6000	$(7, 1) \times 0$		
	$(9, 1) \times 80$	$(9, 1) \times 80$ $(8, 1) \times 20$ $(7, 1) \times 0$	$(9, 1) \times 80$ $(8, 1) \times 20$ $(7, 1) \times 0$
	$(7, 3) \times 8$		
	$(6, 3) \times 8$		
	$(7, 2) \times 3$		
	$(5, 5) \times 1$		
	$(7, 1) \times 0$		

$$\sigma^2 = 0$$

N	F-Y method	proposed method	
		$\varepsilon = CV^*/10$	$\varepsilon = CV^*/20$
3000	$(4, 5) \times 34$	$(7, 1) \times 99$ $(3, 4) \times 1$	$(7, 1) \times 99$ $(3, 4) \times 1$
	$(7, 1) \times 31$		
	$(4, 4) \times 21$		
	$(3, 5) \times 4$		
	$(3, 4) \times 2$		
	$(5, 5) \times 2$		
	$(7, 2) \times 2$		
	$(7, 3) \times 1$		
	$(5, 4) \times 1$		
	$(6, 5) \times 1$		
	$(1, 5) \times 1$		
	$(7, 1) \times 46$		
6000	$(4, 5) \times 39$	$(7, 1) \times 100$	$(7, 1) \times 100$
	$(4, 4) \times 12$		
	$(5, 4) \times 2$		
	$(6, 5) \times 1$		

$$\sigma^2 = 0.01$$

TABLE 4 Frequency distributions of $(\hat{d}_0, \hat{\tau}_0)$ in 100 trials for Model 4, where the bold face represents the correct estimates.

The table shows that the F-Y method gives wrong estimates all the times when $\sigma^2 = 0$, and it estimates the true value only 46 times in 100 trials when $\sigma^2 = 0.01$ even when $N = 6000$. The proposed method gives wrong estimates all the times when $\sigma^2 = 0$, $\varepsilon = CV^*/10$, $CV^*/20$ and $N = 3000$. The results of the proposed method were not changed even when we increased the sample size up to 6000. When $\sigma^2 = 0.01$, it estimates the true value 99 times in 100 trials when $N = 3000$ and it estimates the true value without errors in all combinations of $\varepsilon = CV^*/10$, $CV^*/20$ and $N = 6000$.

5. Application to pulsation data

Tsuda, Tahara and Iwanaga (1992) investigated the pulse at the finger tip capillary vessels obtained from healthy individuals and patients suffering Alzheimer's disease. Embedding the data into 4 dimensional space, they found that the pulsation data from healthy individuals generates chaotic series, but those from the patients suffering from Alzheimer's disease did not. However in their study the dimension of the embedding space was found by trial and errors without regard to the dynamic noise which seems commonly observed in physiological data.

We applied the proposed method to a set of similar data which was also measured at the finger tip capillary vessels from healthy individuals in the rate of 200 points/second. From the study of Tsuda, Tahara and Iwanaga it was assessed that d_0 was around 3 and we supposed $\sigma^2 \leq 0.01$. If this is the case the simulation in the previous section suggests $N \geq 4000$ and $\varepsilon = \frac{CV^*}{20}$. The estimated embedding dimension and delay time when $N = 4000$ and $\varepsilon = \frac{CV^*}{20}$ were $\hat{d}_0 = 2$ and $\hat{\tau}_0 = 1$. To make sure these estimates we also applied the method when $N = 5000$ and $\varepsilon = \frac{CV^*}{10}$. The results are listed in Table 5.

N	$\varepsilon = CV^*/10$	$\varepsilon = CV^*/20$
4000	(2,1)	(2,1)
5000	(2,1)	(2,1)

TABLE 5 The results of application of our method to pulsation data.

The table shows that $\hat{d}_0 = 2$ and $\hat{\tau}_0 = 1$ in all cases. The result differs from Tsuda, Tahara and Iwanaga who estimated $\hat{d}_0 = 3$ and $\hat{\tau}_0 = 25$. The large difference of the estimated delay time is partly due to the difference of the aim: Tsuda, Tahara and Iwanaga selected the delay time aiming to visualize the dynamics, but we estimated it from the view point of prediction. Furthermore, it might due to the characteristics of the delay time which may be described as follows: Suppose for simplicity that $Y_t = F(Y_{t-1})$ when the pulse is measured in the rate of 100 points/second, then it corresponds to $Y_t = F(Y_{t-2})$ when it is measured in the rate of 200 points/second, namely, the delay time do depend on the rate of the measurements. Note that the true delay times are estimated by our method in the simulation.

6. Discussions

It is found that the CV criterion employed in Fueda and Yanagawa (2001) is sensitive to initial values if F is a chaotic map. Several rules are introduced based on the intervals with length ε in this paper to accommodate the sensitivity. The selection of ε is important, but we have no answer for its optimal selection. We used $\varepsilon = CV^*/10$ and $\varepsilon = CV^*/20$ in the simulation and showed that the method works fairly well when $d \leq 4$, σ^2 is small, and $N \geq 1000$.

An advantage of the autoregressive approach is in the sample size requirement; the method may estimate the embedding dimension with substantially less sample size than conventional approaches in physical science, as was discussed in Chen and Tong (1994). Even with this advantage it was shown that $N = 6000$ is not enough when $d = 7$. Note

that the method works even in this case when $\sigma^2 > 0$ since it increases variability. The method fails unless the variance of the dynamic noise is remarkably small. This might appear to be fatal, but not at all; it is shown in the simulation that the method is applicable to data with $\sigma^2 = 0$; that is, the data from deterministic chaos. It could be another, but more theoretical approach in estimating the embedding dimension and delay time than those methods developed in the literatures in physical science.

Acknowledgement

The authors thank the Institute of Statistical Mathematics for the facilities and the use of SGI2800. The authors also wish to express their hearty thanks to the referee for his kind comments.

References

- Abarbanel, H. D. I. (1995). *Analysis of Observed Chaotic Data*, Springer, Berlin.
- Cheng, B. and Tong, H. (1994). Orthogonal projection, embedding dimension and sample size in chaotic time series from a statistical perspective, In Tong, H., editor, *Chaos and Forecasting*, World Scientific, Singapore, London.
- Frazer, J. D. and Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information, *Physical Review A*, **33**, 1134-1140.
- Fueda, K. and Yanagawa, T. (2001). Estimating the embedding dimension and delay time from chaotic time series with dynamic noise, *Journal of the Japan Statistical Society*, **31**, No.1, 27-38.
- Grassberger, P. and Procaccia, I. (1983). Characterization of strange attractors, *Phys. Rev. Letters*, **50**, 346-349.
- Henon, M. (1976). A Two-dimensional Mapping with a Strange Attractor, *Communications in Mathematical Physics*, 69-77.
- Kennel, M. B., Brown, R., and Abarbanel, H. D. I. (1992). Determining minimum embedding dimension using a geometrical construction, *Physical Review*, **45**, 3403-3411.
- Nadaraya, E. A. (1964). On estimating regression, *Theory Prob. Appl.*, **9**, 141-142.
- Tsuda, I., Tahara, T. and Iwanaga, H. (1992). Capillary chaos and its significance in information processing. (in Japanese), *Holonics*, **3**, No.4, 163-176.
- Yonemoto, K. and Yanagawa, T. (1998). Graphical representation of observed chaotic time series and its application. In *proceedings of chaos and non-linear time series symposium* in Hakata.
- Watson, G.S. (1964). Smooth regression analysis, *Sankhyā Ser.A*, **26**, 359-372.

Received December 21, 2001

Revised January 24, 2002